



**HAL**  
open science

# The origins of isolating word structure in eastern Timor

Antoinette Schapper

► **To cite this version:**

Antoinette Schapper. The origins of isolating word structure in eastern Timor. Austronesian Undressed: How and Why Languages Become Isolating, 2020. halshs-02935746

**HAL Id: halshs-02935746**

**<https://shs.hal.science/halshs-02935746v1>**

Submitted on 10 Sep 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The origins of isolating word structure in eastern Timor

Antoinette Schapper  
Lacito-CNRS

This paper addresses the issue of isolating word structure and its origins in the Austronesian and Papuan languages of eastern Timor. McWhorter (2007) claims that both families of languages evidence extensive loss of grammatical complexity as a result of “interrupted transmission” due to significant non-native acquisition. I refute McWhorter’s assertion that the eastern Timor languages are not “normal” through a detailed exposition of their morphological complexities. Whilst recognising that they are isolating leaning, I argue that there is nothing “unnatural” about the grammars of these languages and that phonological changes within the Timorese Sprachbund provide sufficient explanation of their morphological profiles.

Keywords: Timor languages, phonological erosion, irregularity, lexicalisation, isolating word structure, convergence

## 1. Introduction<sup>1</sup>

McWhorter (2007) attempts an exciting piece of research, putting forth the radical argument that all cases where a language appears to have been simplified to an extent not explainable by means of regular linguistic change are due to the intervention of non-native learners. Claims about the apparent lack of complexity in some languages have to be taken with several grains of salt, since McWhorter’s (2007) criteria for assessing complexity are of highly debatable heuristic value (see, e.g., papers in Sampson, Gil, and Trudgill 2009 for an assessment of different claims, among many others). More significant for this paper, however, is the radical claim that particular languages evidence extensive loss of grammatical complexity such that they must be regarded as having had

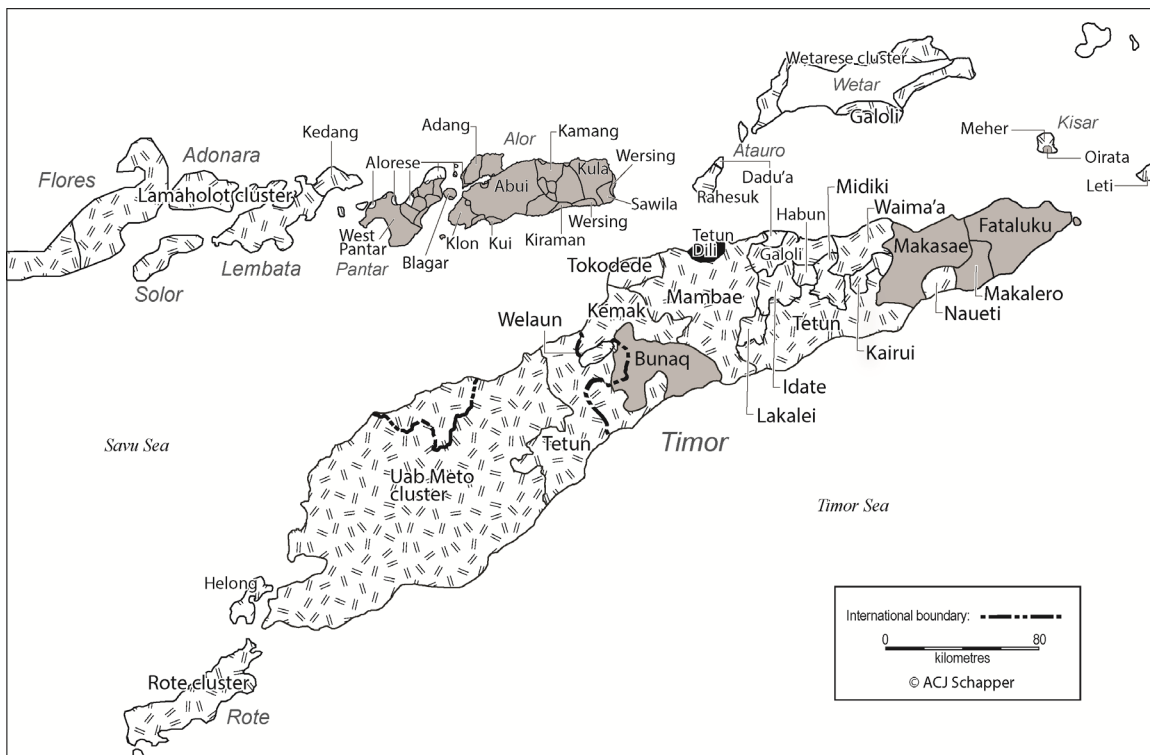
---

<sup>1</sup> This paper was first written in 2009 following my participation in the panel on Isolating Austronesian Languages at ICAL11 in Aussois. Since 2009, I have had the opportunity to learn much more about the Austronesian languages of Timor, together with Rachel Hendery during fieldwork on Tokodede, in supervising Alexandre Veloso’s thesis work on Naueti, and in teaching a Masters course on Waima’a using materials supplied by Nikolaus Himmelmann. Thank you all very much for these opportunities! This paper has also benefited greatly from conversations with Chuck Grimes and data supplied by him from his extensive fieldwork in Timor. I have also learnt a huge amount about the Eastern Timor Papuan languages from Juliette Huber and Aone van Engelenhoven. Thanks to these people, this paper has been greatly expanded and improved since its first drafting. Research funding has come from the Netherlands Organisation for Scientific Research VENI project “The evolution of the lexicon. Explorations in lexical stability, semantic shift and borrowing in a Papuan language family”, the Volkswagen Stiftung DoBeS project “Aru languages documentation”, and the Australian Research Council project (ARC, DP180100893) “Waves of words”. All errors are my own.

their normal development “interrupted” by widespread non-native acquisition. I shall focus on McWhorter’s claims about the languages of eastern Timor that are at odds with actual language data. I argue that the claims reveal a lack of concern for explaining patterns in language over asserting a model.

This paper addresses the issue of isolating word structure and its origins in the Austronesian and Papuan languages of eastern Timor. The island of Timor is located at the eastern end of the Minor Sundic island chain (Map 1). It is home to around a dozen languages of the Austronesian family (speckled on the map) and four Papuan languages of the Timor-Alor-Pantar family (dark grey on the map). Both groups of languages of the eastern half of Timor are largely isolating, typically having little inflectional morphology beyond a set of verbal agreement affixes and a set of possessor affixes. I look at claims made in McWhorter (2007, repeated in 2008) that the isolating structure of the languages of eastern Timor evidence their reduced complexity as the result of significant non-native acquisition in the past. I argue that the evidence for McWhorter’s scenario of ‘interrupted transmission’ causing morphological ‘stripping’ and overall reduced complexity is weak, and does not stand up to scrutiny on proper examination of data from the languages of eastern Timor. I show that these languages evidence the kinds of complexity that McWhorter says to be typical of older languages. In fact, I propose that morphological loss can be accounted for by patterns of phonological changes shared across eastern Timor languages.

Map 1. The languages of eastern Timor and surrounds



This paper is structured as follows. Section 2 outlines McWhorter's view of complexity, while Section 3 presents his arguments for viewing eastern Timor languages as having wide non-native acquisition in their histories. I take a deep dive into data from the Austronesian and Papuan languages of eastern Timor in Sections 4 and 5 respectively, arguing that there is more morphological complexity to be observed in the languages than McWhorter would have us believe. Section 6 then compares the eastern Timor languages to their nearest neighbours and relatives, with a view to understanding how reduced the languages in fact are. I argue that whilst eastern Timor languages show morphological reduction in comparison to their nearest relatives, they cannot be characterised as drastically morphologically stripped and their differences are readily explainable by means of normal processes of language change. In Section 6, I discuss alternative explanations for the isolating structure, both in terms of a Timorese Sprachbund, and a larger, more ancient convergence pattern.

## 2. McWhorter's complexity

McWhorter argues in a series of publications (2001, 2005, 2007, 2008) that isolating word structure in language is 'unnatural'. McWhorter (2001) writes, "[i]n the uninterrupted transmission of a human language, radical loss of complexity throughout the grammar is neither normal, occasional, nor rare, but impossible...". He continues, '[o]lder languages at all times retain a degree of accreted complexity distinguishing them from languages that were born as pidgins'. That is, he claims that isolating languages arise exclusively as the result of 'interrupted language transmission', either from creolization (2001, 2005), or from 'Non-hybrid Conventionalized Second-Language varieties' (2007).

McWhorter (2007) elaborates further on his earlier claims that all cases where a language appears to have been simplified to a degree not explainable by means of regular linguistic change are due to the intervention of non-native learners. He defines a matrix of complexity by which the simplicity of a language can be assessed. Three factors contribute to a language's complexity:

- i. *Overspecification*: This refers to the differing degrees to which languages overtly and obligatorily mark semantic distinctions (McWhorter 2007: 21-29). For McWhorter, overspecification means that languages evidence features such as noun class marking including numeral classifiers, possessive classes such as inalienable versus alienable distinctions, definiteness marking, TAME marking, markers of valency change, multiple degrees of demonstrative gradation, numerous negators expressing different negative semantics, abundant pragmatic particles, etc.
- ii. *Structural elaboration*: This refers to the number of rules in morphosyntax and elements in (morpho)phonology that derive surface structures (McWhorter 2007: 29-33). Linguistic features that McWhorter views as structurally elaborate are, for instance, complex morphophonemics, large phonemic inventories or tones with multiple contrasting levels, grammatical gender systems, and declension and conjugation classes.

- iii. *Irregularity*: This refers to the lexical specification of grammatical features and of paradigm cells that are not the expression of generalised rules, but must be learned by rote. For McWhorter, irregularity in language can be seen in the presence of features like assignment of grammatical gender/noun classes, irregular plural formation, suppletion in inflectional paradigms, etc.

There are, of course, many other ways in which a language could be seen to display complexity, or a lack thereof. McWhorter (2007: 268) gives the example of Pirahã, an indigenous language of South America, observing that the absence of numerals, colour terms, and clausal embedding in the language could be regarded as evidence of its simplicity. Yet, for McWhorter, these features are irrelevant to complexity as he defines it and are merely the result of an “unelaborated cultural perspective among its speakers”. McWhorter sees that despite a small phonemic inventory, Pirahã has ample complexity with its two-tone contrast, inflections for aspect and evidentiality, and nominalisation morphology.

McWhorter maintains that extreme lack of complexity – in the form of overspecification, structural elaboration, and irregularity – is not attributable to chance in the world’s languages. In his view, non-native acquisition tends to shave away features such as these as they are less necessary to communication. According to McWhorter, languages that display high levels of the above features are older languages; while languages that do not have these features have had ‘interrupted language transmission’ as a factor in their history. In the strongest version of his hypothesis McWhorter argues that simple grammars are *impossible* without extensive non-native acquisition.

### **3. McWhorter’s explanation of isolating word structure in Timor languages**

Turning to Austronesian languages, McWhorter (2007: 242-251, 2008) makes the observation that extreme isolating structure appears to be cross-linguistically rare within the Austronesian family. He contrasts the morphologically rich profile of many Austronesian languages to that presented by some of the languages of Flores and of Timor. He explains that these languages show “unusual morphological simplification” (2007: 247), having “shed all or most of their inflections” (2007: 248). He likens them to creoles, stating that it is “extremely unusual for an older language to hover this closely above [the level of complexity of] creoles” (2007: 251), asserting that their isolating structure must be “traced to heavy non-native acquisition at some point in the past” (2007: 248).

On Timor, McWhorter (2007: 242-251, 2008: 175-181) draws a contrast between the many isolating languages of the eastern half of Timor to those spoken in the western half. He writes that Uab Meto (also known as Dawan(ese), Atoni or Timorese) has two sets of subject prefixes with unpredictable distributions, metathesis and irregular verbs, ‘complex’ features which he says to be not present in eastern Timor. Similarly,

McWhorter also notes that Rotinese<sup>2</sup> is of greater complexity than the Austronesian languages of eastern Timor, with verbal subject prefixes, eight numeral classifiers and possessive enclitics conditioned by constituent class. He concludes that western Timor languages show the ‘normal complexity’ (2008: 178) of eastern Austronesian languages that he establishes from inspection of surrounding languages such as Kambara (Sumba), Sika (Flores) or Tukang Besi (south-east Sulawesi).

McWhorter observes that it is not only the Austronesian languages of eastern Timor but also the Papuan ones that are ‘morphologically stripped’ and ‘simplified’ (2007: 248). McWhorter further asserts that the Timor languages contrast with the typically synthetic grammars of Papuan languages in general (McWhorter 2007: 248-249) or their nearest Papuan relatives on New Guinea (McWhorter 2008: 178-179). Although no data is presented in support of either of these claims, McWhorter concludes that the shared simplification of Papuan and Austronesian languages on Timor points to a common event of ‘interrupted transmission’ in their histories. He rejects the idea that the Papuan languages of Timor could have developed in an isolating direction because of contact with their morphologically reduced Austronesian neighbours, declaring that this would mean ascribing a highly unusual degree of structural loss to the contact in question. He admits that areal pressure can result in morphological reduction, but maintains that a mere contact account is not sufficient to explain “why they lost so much morphology overall that they stand as strangely analytic, or analytic-leaning, languages” (McWhorter 2007: 248).

Following Hull (1998, 2001), McWhorter (2007, 2008) hypothesises that the historical event that led to the morphological stripping of Timorese languages was an invasion from Central Maluku approximately eight hundred years ago. This theory is based on a Timorese myth involving incoming Ambonese (Hull 1998: 161-164) and the existence of six Timorese placenames similar to placenames in Ambon (Hull 1998: 162). McWhorter concludes that although the details of the Ambonese invasion are unknown, “treating this migration as the cause of the strangely low level of complexity in Timor languages’ grammar is more scientific than ascribing the anomaly to chance” (2008: 181). Whilst it is not the intention of this article to dissect the claims of Hull which McWhorter bases his reasoning on, it is worth noting that no linguists working in the area have taken them up, let alone even bothered to refute them in writing. Oral traditions such as origin myths are well known in the region to reflect political expediencies rather than historical realities (Wellfelt 2016), while similar placenames are often found over areas where related languages are found due to similarities in the strategies used to name places.<sup>3</sup> What is more, Hull (1998) makes clear that the “Ambonese signature” is

---

<sup>2</sup> Note that there is no single language of Rote, but a cluster of languages, potentially not very closely related to one another (Edwards 2018a, 2018b). “Rotinese” is typically used by McWhorter and in the general literature for the Termanu language as described in the pioneering work of Jonker (1915). I will follow this practice here for the sake of simplicity.

<sup>3</sup> For example, Lutur is a placename found in the Aru, Kei and Tanimbar islands, but the placename itself is not the result of contact between groups on these islands. *Lutur* is a noun

strongest in the languages of Roti and Uab Meto, spoken in western Timor and precisely the groups whose languages McWhorter claims to have normal levels of complexity. In short, the historical scenario used by McWhorter to explain isolating structure in eastern Timor is without merit and I won't engage with it further in this paper.

In what follows, I restrict myself to addressing McWhorter's claim that the languages of Timor are 'morphologically stripped' and, in general, lack the complexity of older languages. My discussion of complexity will focus on morphological complexity, though the reader should bear in mind that morphological complexity is only one kind of complexity. McWhorter (2007) himself emphasises this, but also observes that morphological complexity is still the most decisive feature as it constitutes the first wave of grammatical simplification in the wake of which various complexities of other kinds can remain.

#### 4. Austronesian languages in Timor<sup>4</sup>

The Austronesian languages of Timor are thought to be divided into two subgroups, with some differences between authors (compare Hull 1998 and Edwards 2018a). The Central Timor subgroup is small, containing just Tokodede, Kemak, Mambae and Welaun (Edwards 2019: 42-49). The remaining languages of both east and west Timor, with the exception of Helong, appear to all belong together in a single, large "Timor-Wetar-Babar" subgroup that stretches well beyond Timor to the Babar islands in southern Maluku (see the following for various subgrouping arguments for the region: van Engelenhoven 1987, 2009a, 2010; Mills 1991; Taber 1993; Hull 1998; Edwards 2018a: 86-88). Within this group on Timor, the so-called Kawaimina (*Kairui*, *Waima'a*, *Midiki* and *Naueti*) languages all are very closely related and seem to have arisen out of a differentiated dialect chain. The low-level subgroupings of other Austronesian languages in eastern Timor have not as yet been established, while those of the western Timor languages have been explored extensively in recent times (e.g., Edwards 2018b).

In what follows, I present a wide-range of data from the Austronesian languages spoken in the eastern half of Timor (including the Central Timor languages), illustrating the diverse range of morphological structures that the languages present.

---

meaning 'fort, stone wall' in the indigenous languages of southern Maluku (Schapper 2019) and places with prominent features such as forts were often named after those landmarks.

<sup>4</sup> The following sources were used for the eastern Timor Austronesian languages discussed here: Dadu'a, Penn (2006); Galoli (aka Galolen), Hull (2003); Habun, Hull (2001); Raheuk (aka Hresuk), Boarcceach (2013); Idate, Alcantara (2015); Kemak Atsabe, own fieldnotes, Schapper (2009); Kemak Marobo, Chuck Grimes p.c.; Lakalei, Hull (2001); Mambae Ainaro, own fieldnotes; Mambae Same, Grimes et al. (2014); Naueti, Veloso (2016); Tetun Fehan (West Timor), van Klinken (1999); Tetun Terik (East Timor), Hull (2001); Waima'a, Bowden et. al. (nd.), Hull (2002); Welaun (aka Wekais), da Silva (2012), Edwards (2019).

#### 4.1 Verbal agreement prefixes

As is common in eastern Indonesia, the majority of Austronesian languages in eastern Timor have verbal prefixes agreeing with their subject. Only the Kawaimina languages Tokodede and Kemak are exceptions, having no known verbal agreement prefixes. Examples of these prefixes are given in Table (1). We see that languages differ in the number of persons that are marked. For most of the languages, these prefixes are only found on vowel- and sometimes h-initial roots (often replacing initial *h*, Hull 2001:153-154).<sup>5</sup> In Habun, however, the prefixes occur both on vowel and consonant initial roots. In Mambae Ainaro *n-* appears erratically on a small number of vowel-initial verbs.

Table 1: Subject agreement prefixes in the AN languages of eastern Timor

	Galoli	Habun	Welaun	Tetun Fehan	Lakalei	Mambae Ainaro
1SG	<i>ʔ-</i>	<i>k-</i>	<i>k-</i>	<i>k-</i>	--	--
2SG	<i>m-</i>	<i>m-</i>	<i>m-</i>	<i>m-</i>	<i>m-</i>	--
3SG	<i>n-</i>	<i>n-</i>	<i>n-</i>	<i>n-</i>	<i>n-</i>	<i>n-</i>
1PL.INCL	<i>t-</i>	<i>t-</i>	--	--	--	--
1PL.EXCL	<i>r-</i>	<i>h-/--</i>	--	--	--	--
2PL	<i>r-</i>	<i>h-/--</i>	--	--	--	--
3PL	<i>r-</i>	<i>r-</i>	<i>n-</i>	<i>n-/r-</i>	--	--

Whilst the Austronesian languages above have just a single set of prefixes, there are others that have multiple sets. For example, Idate has one set that occurs on vowel initial verbs and another on consonant initial verbs (Table 2).

Table 2: Idate agreement prefixes paradigms

	Set 1		Set 2	
		‘buy’		‘open’
1SG	--	<i>ala</i>	<i>u-</i>	<i>uloʔe</i>
2SG	<i>m-</i>	<i>mala</i>	<i>o-</i>	<i>oloʔe</i>
3SG	<i>n-</i>	<i>nala</i>	<i>na-</i>	<i>naloʔe</i>
1PL.INCL	--	<i>ala</i>	<i>ta-</i>	<i>taloʔe</i>
1PL.EXCL	--	<i>ala</i>	--	<i>loʔe</i>
2PL	--	<i>ala</i>	--	<i>loʔe</i>
3PL	<i>r-</i>	<i>rala</i>	<i>ra-</i>	<i>raloʔe</i>

Rahesuk also has two sets the forms of which are phonologically conditioned (Table 3). Set 1 goes on verbs with an initial sonorant consonant. The full set appears on initial verbs with initial liquids, but is reduced in different ways on verbs with initial nasals: on

<sup>5</sup> The appearance of agreement prefixes may also be affected by discourse pragmatics in multi-verb clauses (see, e.g., the description of Tetun Fehan subject agreement in van Klinken 1999: 174-176).



*n*-initial verbs the 3<sup>rd</sup> person singular prefix *n*- is lost; on *m*-initial stems the 3<sup>rd</sup> person singular is infixes as <*n*>, while the *m*- prefix for 2<sup>nd</sup> persons and 1<sup>st</sup> person plural exclusive is lost; finally, on *ŋ*-initial stems only the *k*- prefix for 1<sup>st</sup> person plural inclusive and third person plural is retained. Set 2 prefixes appear on vowel initial verbs and are used as infixes on *h*-initial verbs. Verbs beginning with other consonants do not take agreement markers.

Table 3: Rahesuk agreement prefixes paradigms

Set 1				
	L-initial	‘go’	n-initial	‘sow’
1SG	--	<i>laʔa</i>	--	<i>naho</i>
2SG	<i>m</i> -	<i>m</i> <i>laʔa</i>	<i>m</i> -	<i>m</i> <i>naho</i>
3SG	<i>n</i> -	<i>n</i> <i>laʔa</i>	--	<i>naho</i>
1PL.INCL	<i>k</i> -	<i>k</i> <i>laʔa</i>	<i>k</i> -	<i>k</i> <i>naho</i>
1PL.EXCL	<i>m</i> -	<i>m</i> <i>laʔa</i>	<i>m</i> -	<i>m</i> <i>naho</i>
2PL	<i>m</i> -	<i>m</i> <i>laʔa</i>	<i>m</i> -	<i>m</i> <i>naho</i>
3PL	<i>k</i> -	<i>k</i> <i>laʔa</i>	<i>k</i> -	<i>k</i> <i>naho</i>
	m-initial	‘come’	ŋ-initial	‘swim’
1SG	--	<i>ma</i>	--	<i>ŋaŋi</i>
2SG	--	<i>ma</i>	--	<i>ŋaŋi</i>
3SG	< <i>n</i> >	<i>m</i> <i>n</i> <i>a</i>	--	<i>ŋaŋi</i>
1PL.INCL	<i>k</i> -	<i>k</i> <i>m</i> <i>a</i>	<i>k</i> -	<i>k</i> <i>ŋaŋi</i>
1PL.EXCL	--	<i>ma</i>	--	<i>ŋaŋi</i>
2PL	--	<i>ma</i>	--	<i>ŋaŋi</i>
3PL	<i>k</i> -	<i>k</i> <i>m</i> <i>a</i>	<i>k</i> -	<i>k</i> <i>ŋaŋi</i>
Set 2				
	V-initial	‘drink’	h-initial	‘(re)turn’
1SG	--	<i>enum</i>	--	<i>hali</i>
2SG	<i>m</i> -	<i>m</i> <i>enum</i>	< <i>m</i> >	<i>h</i> <i>m</i> <i>ali</i>
3SG	<i>n</i> -	<i>n</i> <i>enum</i>	< <i>n</i> >	<i>h</i> <i>n</i> <i>ali</i>
1PL.INCL	<i>r</i> -	<i>r</i> <i>enum</i>	< <i>r</i> >	<i>h</i> <i>r</i> <i>ali</i>
1PL.EXCL	<i>m</i> -	<i>m</i> <i>enum</i>	< <i>m</i> >	<i>h</i> <i>m</i> <i>ali</i>
2PL	<i>m</i> -	<i>m</i> <i>enum</i>	< <i>m</i> >	<i>h</i> <i>m</i> <i>ali</i>
3PL	<i>r</i> -	<i>r</i> <i>enum</i>	< <i>r</i> >	<i>h</i> <i>r</i> <i>ali</i>

Dadu'a is similarly complex in that there are two agreement sets but verbs are lexically assigned to them (forms and examples in Table 4). In addition, prefixation is associated with a raft of morphophonological changes to roots. For instance, verbs with initial /b/ and /p/ show the following changes when set 1 prefixes are attached: *k* + *b* > *f* and *h* + *p*/*b* > *f*. In addition, Dadu'a also has a set of verbs with irregular prefixal paradigms, some of which are illustrated in Table 5.

Table 4: Dadu'a agreement prefixes paradigms

	Set 1			Set 2		
		'live'	'close'		'blow'	'make'
1SG	--	<i>mia</i>	<i>paʔa</i>	--	<i>afuu</i>	<i>oi</i>
2SG	--	<i>mia</i>	<i>paʔa</i>	<i>m-</i>	<b><i>mafuu</i></b>	<b><i>moi</i></b>
3SG	--	<i>mia</i>	<i>paʔa</i>	<i>n-</i>	<b><i>nafuu</i></b>	<b><i>noi</i></b>
1PL.INCL	<i>k-</i>	<b><i>kmia</i></b>	<b><i>kbaʔa</i></b>	<i>t-</i>	<b><i>tafuu</i></b>	<b><i>toi</i></b>
1PL.EXCL	<i>h-</i>	<b><i>hmia</i></b>	<b><i>faʔa</i></b>	<i>r-</i>	<b><i>rafuu</i></b>	<b><i>roi</i></b>
2PL	<i>h-</i>	<b><i>hmia</i></b>	<b><i>faʔa</i></b>	<i>r-</i>	<b><i>rafuu</i></b>	<b><i>roi</i></b>
3PL	<i>h-</i>	<b><i>hmia</i></b>	<b><i>faʔa</i></b>	<i>r-</i>	<b><i>rafuu</i></b>	<b><i>roi</i></b>

Table 5: Dadu'a irregular prefixing verb paradigms (irregular forms bolded)

	'die'	'go'	'cut'	'enter'	'injure'	'reach'
1SG	<i>mate</i>	<i>laa</i>	<i>looh</i>	<i>tama</i>	<i>namani</i>	<i>raik</i>
2SG	<i>mate</i>	<i>laa</i>	<b><i>hlooh</i></b>	<i>tama</i>	<i>namani</i>	<i>raik</i>
3SG	<b><i>nate</i></b>	<i>laa</i>	<i>looh</i>	<i>tama</i>	<i>namani</i>	<i>raik</i>
1PL.INCL	<b><i>kmate</i></b>	<b><i>kaʔa</i></b>	<b><i>klooh</i></b>	<i>tama</i>	<b><i>tamani</i></b>	<b><i>rai</i></b>
1PL.EXCL	<b><i>hmate</i></b>	<b><i>hlaa</i></b>	<b><i>hlooh</i></b>	<b><i>tahma</i></b>	<b><i>ramani</i></b>	<b><i>hraiik</i></b>
2PL	<b><i>hmate</i></b>	<b><i>hlaa</i></b>	<b><i>hlooh</i></b>	<b><i>tahma</i></b>	<b><i>ramani</i></b>	<b><i>hraiik</i></b>
3PL	<b><i>hmate</i></b>	<b><i>hlaa</i></b>	<b><i>hlooh</i></b>	<b><i>tahma</i></b>	<b><i>ramani</i></b>	<b><i>hraiik</i></b>

Far from being stripped of verbal inflectional morphology, we have seen that most Austronesian languages of eastern Timor have subject agreement prefixes that appear on at least a subset of the verbal lexicon. For several languages we find multiple sets of verbal agreement prefixes whose choice of host may be lexical or phonological. Morphophonemic processes also can be observed to frequently play a role in determining the surface forms of prefixes and roots.

#### 4.2 Derivational prefixes and associated complexification

Austronesian languages are well-known for their derivational morphology (see, e.g., Blust 2014). Like other eastern Austronesian languages, however, the languages of Timor have none of the voice morphology that characterises languages in western Indonesia and the Philippines. Nonetheless, a range of Austronesian derivational prefixes is still found in eastern Timor languages. Depending on the language, these prefixes may be fossilised, productive or have, in some cases, even fused with inflectional morphology. They have not been simply 'shed', but are accreted as part of the system. What is more, accreted prefixes can also be observed to have caused considerable complexification in other domains of linguistic structure.

A prime example of this is the large consonant inventories of Waima'a and Naueti that have come into being through the fossilisation of prefixes in the languages. Considerably above the average Timorese language consonant inventory of 12-15 consonants, Naueti has 27 native consonant phonemes (Table 6) and Waima'a 30 (Table 7). Beyond their

large size, these consonant phoneme inventories are cross-linguistically unusual in that they include phonological rarities such as ejectives, preglottalised consonants, postglottalised consonants, contrastively aspirated plosives, and voiceless sonorants. The historical source of these additional consonant phoneme series in prefixes is apparent from their being limited to word-initial position and being present in relatively few items compared to their “regular” (non-glottalised, non-aspirated, non-devoiced) counterparts.

Table 6: Naueti consonant phonemes

		Labio- velar	Bilabial	Alveolar	Velar	Glottal
Plosive	voiceless			t	k	ʔ
	voiced		b	d	g	
	aspirated		p <sup>h</sup>	t <sup>h</sup>	k <sup>h</sup>	
Fricative	voiceless			s		h
Nasal	voiceless		m̥	n̥		
	voiced		m	n		
	preglottalised		ʔm	ʔn		
Lateral	voiceless			l̥		
	voiced			l		
	preglottalised			ʔl		
Rhotic	voiceless			r̥		
	voiced			r		
	preglottalised			ʔr		
Approximant	voiceless	w̥				
	voiced	w		j		
	preglottalised	ʔw				

Table 7: Waima'a consonant phonemes

		Labial	(Post)- Alveolar	Velar	Glottal
Plosive	voiceless		t	k	ʔ
	voiced	b	d	g	
	aspirated	p <sup>h</sup>	t <sup>h</sup>	k <sup>h</sup>	
	ejective	pʔ	tʔ	kʔ	
Fricative	voiceless		s		h
	glottalised		sʔ		
Nasal	voiceless	m̥	n̥		
	voiced	m	n		
	glottalised	mʔ	nʔ		
Lateral	voiceless	l̥			
	voiced	l			
	glottalised	lʔ			
Rhotic	voiced	r			
	glottalised	rʔ			
Approximant	voiceless	w̥			
	voiced	w			
	glottalised	wʔ			

Waima'a and Naueti aspirated stops and voiceless sonorants have their origins in the absorption of a prefix through the following steps: PMP \*pa-<sup>6</sup> > \*ha- (cf. Tetun *ha-*) > \*h- > \*hC > C<sup>h</sup> / \_plosive, C̥ / \_sonorant, as exemplified in (1) and (2). Once having entered the phonemic inventories of the languages, these phonological features appear to have taken on a life of their own, assimilating additional lexemes that were never marked by \*pa-.<sup>7</sup> What's more, they have been extended to create new phonemes, notably, /p<sup>h</sup>/ in Waima'a and Naueti appears to have developed analogously to /k<sup>h</sup>/ and /t<sup>h</sup>/ for the assimilation of Tetun (and other) loans with initial f.

#### Origin of Waima'a-Naueti aspirated plosives

- (1) PMP \*pa- + \*kaən 'eat' > Waima'a, Naueti *k<sup>h</sup>aa* (cf. PMP \*k > Waima'a-Naueti k, e.g., PMP \*kahiw 'tree, wood' > Waima'a, Naueti *kai* 'tree, wood')

<sup>6</sup> Note that the exact source prefix cannot be regarded as certain at this stage. I associate the origin of the Kawaimina aspirated plosives and voiceless sonorants here with the PMP \*pa- which was a causative prefix. However, the lack of causative semantics on many of the relevant forms and cognate causative morphemes on many of the relevant forms in neighbouring languages suggests that another morpheme may also be involved. A reviewer suggests that a fossilised agreement prefix \*h- such as that found in Habun an Dadu'a would be a possibility.

<sup>7</sup> Veloso (2016: 4) notes that in general across Kawaimina languages there is a significant lack of correspondence in aspirated stops and voiceless sonorants.

PMP \*pa- + \*takut ‘fear’ > Waima'a, Naueti *tʰaku* ‘fear’ (cf. Tetun *hamtaʔuk* ‘be in fear (of something)’, *taʔuk* ‘be fear’) (cf. PMP \*t > Waima'a-Naueti t, e.g., PMP \*tasik ‘sea’ > Waima'a, Naueti *tasi* ‘sea’)

Origin of Waima'a-Naueti voiceless sonorants

- (2) PMP \*pa- + \*bahuq ‘odour, stench’ > Waima'a *ʔau*, Naueti *ʔou* ‘stink’ (cf. PMP \*b > Waima'a-Naueti w, e.g., PMP \*buaq ‘fruit’ > Waima'a *wuo*, Naueti *wua* ‘fruit’)

PMP \*pa- + \*ma-hataq ‘raw, uncooked’ > Waima'a, Naueti *ʔata* ‘raw, uncooked, unripe, green’ (cf. PMP \*m > Waima'a-Naueti m, e.g., PMP \*manuk ‘chicken’ > Waima'a, Naueti *manu* ‘bird, chicken’)

PMP \*pa- + \*laRiw ‘run, run away’ > Waima'a *ʔai* ‘quick’, Naueti *ʔai~ʔai* ‘very fast, immediate’ (cf. Kemak Atsabe *plai*, Dadu’a *hlai*, Tetun Terik *halai* ‘run’ < \*pa-laRiw) (cf. PMP \*l > Waima'a-Naueti l, e.g., PMP \*qalima ‘hand, five’ > Waima'a, Naueti *lima* ‘arm, hand’)

A similar process of prefix absorption can be seen to have given rise to Naueti Waima'a glottalised consonants through the following steps: PMP \*ka-<sup>8</sup> > \*k- > \*ʔ- > ʔC (Naueti, where C is [+sonorant]) > Cʔ (Waima'a, where C is [-plosive]). I present some examples of the emergence of glottalised sonorants in Waima'a-Naueti in (3). Thus far, there are no PMP etyma reflected in Waima'a that have either /sʔ/ or a member of the ejective phoneme series (a known phonetic progression of post-glottalisation). In addition to the absence of corresponding glottalised phonemes in Naueti, this indicates again that these additional phonemes were created by extension, particularly for the assimilation of borrowings.

Origin of Waima'a-Naueti glottalised consonants

- (3) PMP \*ka- + \*nahik ‘climb’ > Naueti *ʔnai*, Waima'a *nʔai* ‘climb’ (cf. PMP \*n > Waima'a-Naueti n, e.g., PMP \*nunuk ‘banyan tree’ > Waima'a (*kai-*)*nunu*, Naueti *nunu*)

PMP \*ka- + \*muRmuR ‘gargle, rinse the mouth’ > Naueti *ʔmumu* ‘hold between teeth’, Waima'a *mʔumu* ‘rinse’ (cf. Tetun *hak-mumu* ‘to wash or rinse the mouth, to gargle’) (cf. PMP \*m > Waima'a-Naueti m, e.g., PMP \*matay ‘die, dead’ > Waima'a, Naueti *mata*)

PMP \*ka- + \*waRi ‘sun’ > Naueti *ʔwai* ‘dry in sun’, Waima'a *wʔai* ‘dry in sun’ (cf. PMP \*w > Waima'a-Naueti w, e.g., PMP \*wahiR ‘water’ > Waima'a, Naueti *wai*)

Whilst the same level of phonological complexification that we observe in Waima'a and Naueti is not found elsewhere, PMP verbal morphology is not simply shed in other

---

<sup>8</sup> In some cases, this prefix may have been originally been PMP \*paka- that subsequently reduced to \*ka- > \*k- (cf. Tetun *hak-*). As with PMP \*pa- above, it is by no means certain that PMP \*ka was the (only) source for the Kawaimina ejectives and preglottalised stops. Other prefixes or phonological phenomena may have played a role in giving rise to these unusual segments.

Austronesian languages of Timor. Moreover, we do find instances of fossilised morphology giving rise to additional phonemes in other Timorese languages. For instance, *m-* (< PMP \*ma-) is found fossilised on many stative monovalent verbs in many languages of eastern Timor, as illustrated with Kemak Atsabe in (4). PMP \*p is usually reflected as /p/ and PMP \*b as /h/ in Kemak Atsabe. However, \*p > b in Kemak Atsabe under prefixation of \*ma- as follows: \*ma- + \*p > \*mp > \*mb > b (Blust 2008: 96-97). In other situations where consonant clusters are created due to the presence of *m-*, an epenthetic vowel is inserted between /m/ and the first consonant of the root. This unstressed vowel may be realised as schwa or harmonised to the first vowel of the root.

Kemak Atsabe stative prefix

(4)	<i>mdu</i>	‘sit’	[mə'du ~ mu'du]	< PMP *ma- + *tudan
	<i>mnahu</i>	‘fall’	[mə'nahu ~ ma'nahu]	< PMP *ma- + *nabuq
	<i>mnaru</i>	‘long’	[mə'naru ~ ma'naru]	< PMP *ma- + *anaduq
	<i>banasa</i>	‘hot’	[ba'nasa]	< PMP *ma- + *panas

The high frequency with which some prefixes are found suggests that prefixes may have been productive until quite recently. An example of this is the *k-* (< \*ka-, itself perhaps a reduction of PMP \*paka-) prefix which is found on a large number of Tokodede verbs, but not on cognates in nearby languages, as in (5). As in the Kemak examples above, an epenthetic vowel is inserted between /k/ and the first consonant of the root to break up the resulting consonant cluster.

Tokodede *k-* prefix

(5)	<i>kbaas</i>	‘slap’	[kə'baas ~ ka'baas]	< *k-baas	cf. Welaun <i>basa</i> , Kemak <i>basa</i> , Mambae <i>baas</i> , Tetun <i>basa</i> ‘slap’
	<i>kdede</i>	‘knock’	[kə'dede ~ ke'dede]	< *k-dede	cf. Tetun <i>dere</i> ‘hit repeatedly’
	<i>kmus</i>	‘kiss, suck’	[kə'mus ~ ku'mus]	< *k-mus	cf. Kemak Atsabe <i>muusu</i> ‘suck’
	<i>kdula</i>	‘turn’	[kə'dula ~ ku'dula]	< *k-dula	cf. Tetun <i>dulas</i> , <i>hak-dulas</i> , Dadu'a <i>dulah</i> ‘twist, wind’
	<i>kdole</i>	‘crawl’	[kə'dole ~ ko'dole]	< *k-dole	cf. Tetun <i>dolar</i> , Dadu'a <i>dolah</i> ‘crawl’
	<i>kbut</i>	‘close eyes’	[kə'but ~ ku'but]	< *k-but	cf. Welaun <i>buta</i> ‘sleep’

In fact, other eastern Timor Austronesian languages have clearly productive verbal prefixes marking valency changes. Most widespread are causative prefixes (e.g., Welaun *a-* and Tetun *ha-*, reflecting PMP \*pa-). Idate has a fuller range of valency changing suffixes: *si-* (and its allomorph *di-* occurring before /l/) marking an anticausative

derivation of a transitive verb (6); *a-* marking a causative derivation of a intransitive verb, (7), and; *ma-* marking stativity (as opposed to dynamic) on intransitive verbs (8).

Idate valency-changing morphology

- (6) Transitive ~ anticausative alternation
- a. *au u-lo?e lala mata-k*  
 1SG 1SG-open path eye-1SG  
 ‘I open the door (lit. path eye).’
- b. *lala mata-k di-lo?e*  
 path eye-1SG ANTIC-open  
 ‘The door (lit. path eye) opens itself.’ or ‘The door is open.’
- (7) Intransitive ~ causative alternation
- a. *туру, аsu!*  
 descend dog  
 ‘Get down, dog.’
- b. *ами а-туру bandera.*  
 1PL.EXCL CAUS-descend flag  
 ‘We lower the flag.’
- (8) Stative ~ causative alternation
- a. *au ma-nahu hori kareta*  
 1SG STAT-fall from car  
 ‘I fell out of the car.’
- b. *au a-nahu livru*  
 1SG CAUS-fell book  
 ‘I dropped the book.’

Of course, derivational morphology such as this does not have the same status as inflectional morphology in McWhorter’s model (2008: 18-20). But in many eastern Timor Austronesian languages, causatives are not synchronically separable from inflections. Table 8 presents languages in which we find subject agreement prefixes have fused with causative \*pa-. In Waima'a *ra-* (a fusion of a 3PL prefix \*da- with \*pa-) has generalised to all persons; the absence of cognates in its sister language, Naueti, may indicate the Waima'a causative is a borrowing from Galoli.

Table 8: Subject agreement markers fused with causative

	Eastern Tetun†	Dadu'a	Waima'a
1SG	<i>ka-</i>	<i>a-</i>	<i>ra-</i>
2SG	<i>ma-</i>	<i>ma-</i>	
3SG	<i>na-</i>	<i>na-</i>	
1PL.INCL	<i>ra-</i>	<i>ta-</i>	
1PL.EXCL		<i>ra-</i>	
2PL			
3PL			

† These come from Hull (2001:150) Note that van Klinken (1999: 172) does not regard the Tetun Fehan verbal inflections fused with the causative *ha-* prefix, but rather a result of regular morphophonological rule where initial *h* is replaced by an inflection, as described in the previous section. Such a rule would, presumably, explain the origin of the fused forms.

Galoli itself has two series of valency changing morphemes that are fused with subject inflections, as presented in Table 9. The anticausative paradigm is characterised by several morphophonemic rules that determine its surface form (on b-initial verbs, its form is *Ca-*, on g-initial verbs *Cam-*, on s-initial verbs *Can-*, on l-initial verbs *Car-*, and elsewhere *Cak-*).

Table 9: Galoli fused subject agreement and valency changing prefixes

	Causative	Anticausative
1SG	<i>ʔa-</i>	<i>ʔak-</i>
2SG	<i>ma-</i>	<i>mak-</i>
3SG	<i>na-</i>	<i>nak-</i>
1PL.INCL	<i>ta-</i>	<i>tak-</i>
1PL.EXCL	<i>ra-</i>	<i>rak-</i>
2PL		
3PL		

### 4.3 Possessive morphology and possessive classes

Possession in the Austronesian languages of eastern Timor is a domain that exhibits a range of complex structures.

Numerous Austronesian languages of eastern Timor have a paradigm of person-number suffixes occurring on nouns that encode possessors (Table 10). These suffixes represent continuations of conservative Austronesian inflectional morphology, despite some obvious paradigm levelling. Even in languages like Kemak which is lacking verbal prefixes, we find a full paradigm of possessor suffixes. Likewise, Waima'a and Naueti retain the possessive suffix *-n* in the third person, although they have no verbal inflections whatsoever.



Table 10: Possessor suffixes

	Galoli	Kemak Atsabe	Idate	Lakalei	Tetun Fehan	Waima'a	Naueti
1SG	<i>-k</i>	<i>-gV</i>	<i>-k</i>	<i>-k</i>	<i>-n</i>	--	--
2SG	<i>-m/--</i>	<i>-mV</i>	<i>-n</i>			--	--
3SG	<i>-n</i>	<i>-V</i>				<i>-n</i>	<i>-na</i>
1PL.INCL	<i>-r</i>	<i>-rV</i>	<i>-r</i>	<i>-n</i>	<i>-n/-r</i>	--	--
1PL.EXCL						--	--
2PL						--	--
3PL						<i>-n</i>	<i>-na</i>

While one suffixal set is typical, Welaun has two sets of phonologically conditioned possessive suffixes: Set 1 used on nouns with a final vowel, and Set 2 used on nouns with a final consonant (Table 11). Dadu'a is unusual in Timor in that it has a paradigm of possessor prefixes (Table 12).

Table 11: Welaun possessor suffixes

	Set 1	Set 2
1SG	<i>-k</i>	<i>-aak</i>
2SG	<i>-n</i>	<i>-aan</i>
3SG	<i>-n</i>	<i>-aan</i>
1PL.INCL	<i>-t</i>	<i>-aat</i>
1PL.EXCL	<i>-t</i>	<i>-aat</i>
2PL	?	?
3PL	<i>-n</i>	<i>-aan</i>

Table 12: Dadu'a possessor prefixes

1SG	<i>a-</i>
2SG	<i>o-</i>
3SG	<i>ni-</i>
1PL.INCL	<i>ita-</i>
1PL.EXCL	<i>ami- ~ am-</i>
2PL	<i>mi-</i>
3PL	<i>sia- ~ si-</i>

Across eastern Indonesia and Oceania, possessive suffixes are associated with inalienable possession, while unbound possessive markers, typically preceding the possessum, are widely used for alienable possession (Donohue & Schapper 2008). In eastern Timor, such a system of possessive classification is found in Kemak. We see in (9a) that the possessor of the inalienable body part noun *gara-* 'head' is encoded by a possessive suffix, while in (9b) a free possessive pronoun encodes the possessor of the alienable noun *uma* 'house'.

## Kemak Atsabe possessive classes

## (9) Inalienable

- a. *gara-ga*  
head-1SG  
'my head'

## Alienable

- b. *au uma*  
1SG.POSS house  
'my house'

In most Austronesian languages of eastern Timor, however, the morphosyntactic distinction between alienable and inalienable possession such as found in Kemak has broken down. Possessive suffixes and free possessive markers have instead entered into different paradigmatic relationships with one another. In Waima'a and Naueti, the split is

in person: third person possessors of all kinds are encoded with a suffix, e.g., Naueti *-na* (10a and b), while a free possessive marker is used for other persons, such as the Naueti 1<sup>st</sup> person singular possessive form *au* (10c and d).

Naueti possessive coding

- (10) 3<sup>rd</sup> person possessors
- |  |   |
|--|---|
| a. <i>uma-na</i><br>house-3<br>'his/her/their house' | b. <i>lima-na</i><br>arm-3<br>'his/her/their arm' |
|--|---|
- 1<sup>st</sup> person possessors
- |   |  |
|---|--|
| c. <i>au</i> <i>uma</i><br>1SG.POSS house<br>'my house' | d. <i>au</i> <i>lima</i><br>1SG.POSS arm<br>'my arm' |
|---|--|

In Idate, the split between possessive marking strategies is morphophonological: nouns ending in a vowel take possessor suffixes (11a and b), while nouns ending in a consonant take free possessive markers (11c).

Idate possessive coding

- (11) Vowel-final nouns
- |   |  |
|---|--|
| a. <i>namo-k</i><br>garden-1SG<br>'my garden' | b. <i>ibo-k</i><br>mouth-1SG<br>'my mouth' |
|---|--|
- Consonant-final nouns
- |   |
|---|
| c. <i>betuk</i> <i>auk</i><br>bamboo    1SG.POSS<br>'my bamboo' |
|---|

In these Austronesian languages, a remnant of the system of possessive classification is that typical inalienable nouns like body part nouns occur obligatorily with a possessor, while alienable nouns do not require the expression of a possessor to be well-formed. This feature is relatively rare world-wide and represents one manifestation of alienable/inalienable possessive systems that McWhorter deems complex (Bickel & Nichols 2013). Timorese languages add the further complexity of allowing free possessive markers to be either pre-posed or post-posed to the possessed noun, with which fine shades of closeness in possessive relations can be signalled (see van Klinken 1999: 145-152; Schapper 2009).

Possessive morphology is also widely found in the Austronesian languages of eastern Timor in non-possessive contexts, with possessor morphology being used in attributive constructions (cf. Ross 1998 on similar constructions in Oceanic). In Kemak, for instance, nominal (and less often verbal) attributes can be marked as if they were possessed with the 3<sup>rd</sup> person affix *-V*, while the referent noun of the NP behaves like a possessor, occurring before the possessed attribute. Attributive marking with the

possessor suffix is not limited to inalienably possessed nouns, but is found on a wide range of items, such as *tasi* ‘sea’ and *mate* ‘dead’ in (12).

Kemak Atsabe possessive attributive marking

- (12) a. *nipe tasii* ‘sea snake’ < *nipe* ‘snake’ + *tasi-V* ‘sea-3SG’  
 b. *nua matee* ‘old coconut’ < *nua* ‘coconut’ + *mate-V* ‘dead-3SG’

The parallel between possessive and attributive marking is illustrated on the basis of Nauti *-na* in (13) and Idate *-n* (14). Similar constructions have been described for Waima'a (Bowden et al. nd) and Welaun (Edwards 2019: 39-40).

Nauti *-na*

- |      |                         |                          |
|------|-------------------------|--------------------------|
| (13) | Possessive <i>-na</i>   | Attributive <i>-na</i>   |
|      | a. <i>asukai bui-na</i> | b. <i>asukai riku-na</i> |
|      | man cat-3               | man rich-3               |
|      | ‘man’s cat’             | ‘rich man’               |

Idate *-n*

- |      |                      |                           |
|------|----------------------|---------------------------|
| (14) | Possessive <i>-n</i> | Attributive <i>-n</i>     |
|      | a. <i>ni iwa-n</i>   | b. <i>ruut hutu-n isa</i> |
|      | 3SG mouth-3SG        | grass bind-3SG one        |
|      | ‘his, her mouth’     | ‘a bundle of grass’       |
|      |                      | (lit. one bound grass)    |

Again in the domain of possession, the languages of eastern Timor manifest a range of features from inflectional morphology to (in)alienability contrasts that are not consistent with McWhorter’s picture of stunningly ‘simple’ languages.

#### 4.4 Synchronic metathesis

Synchronic metathesis refers to a process whereby the expected linear ordering of sounds in a word is reversed in certain morphosyntactic environments, thus, *xy* becomes *yx*. Metathesis is a striking feature in the Austronesian languages of south-west Maluku, being reported in many of the languages in Timor and southern Maluku (Schapper 2015: 135-138). McWhorter claims that this typologically unusual morphophonological feature is limited to western Timor (found in, e.g., Helong, Bowden 2010; Uab Meto varieties, Edwards 2016). In fact, metathesis is found among the languages that McWhorter points to as being the most simplified in Timor.

Mambae, one of the languages McWhorter names as being most simplified in Timor (2007: 247), metathesis and the related process of apocope is productively used to mark dependency relationships in phrases. In Mambae Ainaro, for example, many nouns have two forms: a vowel-final form which appears phrase finally, and a consonant-final form which appears phrase non-finally, as in the NN compounds in (15). The different forms that are realised in Mambae Ainaro by metathesis are the result of interactions of vowels with one another. Final high vowels such as in (15a) are maintained when they

metathesise into a position next to a low vowel (15b). The mid-vowel /e/ (15c) is assimilated when it metathesises into a position next to /i/ (15d). Similarly, an /a/ vowel (15d) is deleted when it metathesises into a position next to /i/ (15c).

Mambae Ainaro metathesis

- |      |  |   |
|------|--|---|
| (15) | Vowel-final form   | Consonant-final form  |
|      | a. <i>kud tali</i><br>horse rope<br>'bridle' (lit. 'horse rope') | b. <i>tail mata</i><br>rope eye<br>'trap' (lit. 'rope eye') |
|      | Vowel-final form   | Consonant-final form  |
|      | c. <i>an hine</i><br>child female<br>'daughter'                  | d. <i>hiin ana</i><br>female child<br>'girl'                |

For the Same dialect of Mambae, Chuck Grimes (p.c.) estimates metathesis to affect some 30% of the nouns and verbs of the language, as well as some members of closed classes such as pronouns and numerals. A selection of the items from Grimes et al. (2014) is presented in Table 12. Here we see that metathesis results in several morphophonemic changes in the surface form of the metathesised items. Most obvious is the assimilation of final /a/ to the quality of the previous vowel on metathesis. Additionally, on metathesis, final /i/ optionally lowers to /e/ when the preceding vowel is /a/.

Table 12: Metathesising items in Mambae Same

	Vowel-final	Consonant-final		Vowel-final	Consonant-final
IPL.EXCL	<i>ami</i>	<i>aim ~ aem</i>	'mouth'	<i>kuku</i>	<i>kuuk</i>
'when'	<i>arfila</i>	<i>arfiil</i>	'tongue'	<i>lama</i>	<i>laam</i>
'child'	<i>ana</i>	<i>aan</i>	'hand, arm'	<i>lima</i>	<i>liim</i>
'slap'	<i>basa</i>	<i>baas</i>	'front'	<i>muna</i>	<i>muun</i>
'approach'	<i>fedesi</i>	<i>fedeis</i>	'money'	<i>osa</i>	<i>oos</i>
'crow' (v)	<i>foni</i>	<i>foin</i>	'road'	<i>sala</i>	<i>saal</i>
'gather'	<i>futu</i>	<i>fuut</i>	'return'	<i>silā</i>	<i>siil</i>
'stone'	<i>hatu</i>	<i>haut</i>	'grandparent'	<i>tata</i>	<i>taat</i>
'female'	<i>hina</i>	<i>hiin</i>	'year'	<i>tona</i>	<i>toon</i>
'one'	<i>ida</i>	<i>iid</i>	'worm'	<i>ula</i>	<i>uul</i>
IPL.INCL	<i>ita</i>	<i>iit</i>	'house'	<i>uma</i>	<i>uum</i>

Kemak Marobo also has synchronic metathesis, though it appears to be less extensive than that found in Mambae dialects (Chuck Grimes p.c.). (16) presents some preliminary examples of the metathesis. The morphosyntactic rules governing this phenomenon in Kemak remain to be determined by future work. What is clear from these examples is that several morphophonemic rules likely impact on the surface form of the metathesised items, just as in Mambae.

Kemak Marobo metathesis

- |  |  |
|--|--|
| <p>(16) Vowel-final form</p> <p>a. <i>manu hui</i><br/>         chicken wild<br/>         ‘bird’ (lit. ‘wild chicken’)</p> <p>Vowel-final form</p> <p>c. <i>ama naʔi</i><br/>         father royal<br/>         ‘father’s elder brother’</p> | <p>Consonant-final form</p> <p>b. <i>man telo-n ~ maun telo-n</i><br/>         chicken egg-3SG<br/>         ‘chicken egg’</p> <p>Consonant-final form</p> <p>d. <i>aam cuan</i><br/>         father old<br/>         ‘grandfather’</p> |
|--|--|

In sum, synchronic metathesis is attested to be present in what are among the most isolating languages in eastern Timor.<sup>9</sup>

#### 4.5 Numeral agreement

Numeral classifiers are found in the Austronesian languages of eastern Timor, however, in relatively small numbers compared to many other Austronesian languages. Of the languages which have been described, Tetun Fehan is described as having around a dozen different numeral classifiers (van Klinken 1999: 140ff), and at least four numeral classifiers have been identified for Naueti (Veloso 2016: 46). Despite this, there is a feature of numerals in Kawaimina languages that is similar to numeral classification, but is more complex in that it is inflectional.

Simplex numerals between ‘two’ and ‘nine’ in both Waima'a and Naueti must agree in animacy with the referent of the quantified noun. Agreement is indexed by prefixes on the numeral that distinguish HUMAN versus NONHUMAN. The agreement behaviour for numerals is illustrated for the two languages in (17) and (18). In both languages the quantity interrogative ‘how much, many?’ also takes the numeral agreement prefixes to agree with the animacy of the referent whose quantity is questioned.

Waima'a numeral agreement

- |   |  |
|---|--|
| <p>(17) HUMAN agreement</p> <p>a. <i>anu-ata wuo-hitu</i><br/>         woman HUM-seven<br/>         ‘seven women’</p> | <p>NONHUMAN agreement</p> <p>b. <i>kumu kai-hitu</i><br/>         pigeon NHUM-seven<br/>         ‘seven pigeons’</p> |
|---|--|

Naueti numeral agreement

- |   |   |
|---|---|
| <p>(18) HUMAN agreement</p> <p>a. <i>kii wua-lima</i><br/>         person HUM-five<br/>         ‘five people’</p> | <p>NONHUMAN agreement</p> <p>b. <i>uma kai-lima</i><br/>         house NHUM-five<br/>         ‘five houses’</p> |
|---|---|

The Waima'a and Naueti agreement prefixes represent grammaticalisations of numeral classifiers that were independent lexical items: the HUMAN prefix originates in

---

<sup>9</sup> Owen Edwards (p.c.) states that there is also reason to believe that synchronic metathesis is present in Tokodede, a close relative of Mambae and Kemak.

PMP \*buaq ‘fruit’, while the NONHUMAN prefix is from PMP \*kahiw ‘tree, wood’. In both languages, the numeral agreement prefixes are still transparently related to the lexemes from which they grammaticalised (cf. Waima'a *wuo*, Naueti *wua* ‘fruit’, and Waima'a, Naueti *kai* ‘tree, wood’).

A comparative study of numerals in Timor and surrounds suggests that this grammaticalisation did not occur in the immediate ancestor of Waima'a and Naueti (i.e., proto-Kawaimina), but at an earlier stage within the subgroup. Right across languages of the eastern half of the Timor-Babar subgroup, numerals between 2 and 9 occur with fossilised agreement prefixes, reflecting the HUMAN agreement prefix from PMP \*buaq (e.g., numerals 2-5 in Kisar *wo-roʔo*, *wo-kelu*, *wo-ʔakka*, *wo-lima*, and in Dadu'a *wa-rua*, *wa-telu*, *wa-ak*, *wa-lima*).<sup>10</sup> This indicates that, despite exhibiting little person-number inflection on verbs and nouns, Kawaimina languages have not been entirely stripped of inflectional morphology, but preserve agreement prefixes that have lost productivity in nearby related languages.

## 5. Papuan languages of Timor<sup>11</sup>

There are four Papuan languages spoken in two parts of Timor: Bunaq is located in central Timor, while Makasae, Makalero and Fataluku occupy a contiguous region at the island's eastern tip. On Kisar Island, just off the north-eastern end of Timor, is a fifth language, Oirata, a close relative of Fataluku. Bunaq and the four Eastern Timor languages form two primary subgroups of a single language family. Within the Eastern Timor subgroup, Fataluku and Oirata subgroup together, as do Makasae and Makalero. The remaining members of the family are spoken on and between the Alor and Pantar islands, forming a third primary subgroup of Timor-Alor-Pantar.

The nearest relatives of TAP languages, Mbaham, Iha and Kalamang, are spoken on and around the Bomberai peninsula at the western tip of New Guinea. Together, TAP languages and West Bomberai languages have been seen to form the western extreme of the hypothesised Trans-New Guinea family (Usher & Schapper ms).

In the following sections, I illustrate the structures of the Papuan languages of Timor, highlighting areas in which they display indexes of McWhorter's complexity, namely, overspecification, structural elaboration and irregularity.

---

<sup>10</sup> It is not clear that the NONHUMAN agreement prefix from PMP \*kahiw is found outside of Kawaimina languages. The numerals 2-5 in Galoli *i-rua*, *i-telu*, *i-haat*, *i-lima*, for example, still appear to be reflexes of the HUMAN prefix.

<sup>11</sup> The following are the sources for the Papuan languages discussed here: Bunaq, Schapper (2010a), own fieldnotes; Fataluku, van Engelenhoven (2009b), van Engelenhoven & Huber (forthcoming), Heston (2015); Makalero, Huber (2011); Makasae, Huber (2008), Correia (2011); Oirata, de Josselin de Jong (1937).

### 5.1 Person agreement prefixes

A single paradigm of agreement prefixes, occurring on both verbs and nouns, is found in Bunaq (Table 13). The paradigm consists of three person prefixes and two valency reducing prefixes. On prefixation to consonant initial roots, consonantal prefixes appear with an epenthetic vowel. Vowels in prefixes harmonise with the first vowel of the root, similarly to what is found with prefixes in the neighbouring Austronesian languages, Kemak and Tokodede (see Section 4.2). Where the root is vowel initial, the prefixal vowel is deleted. This means that for the first person inclusive and second person, there is no surface manifestation of the prefix (marked as  $\emptyset$ -). On verbs, a prefix typically coindexes an animate P, though there are a small number of verbs with a prefix for animate S. On nouns, a prefix indicates an inalienable possessor.

Table 13: Bunaq agreement prefixes with examples

		<i>bol</i> ‘value’	<i>wit</i> ‘fetch’	<i>il</i> ‘water’	<i>obon</i> ‘hang’
1EXCL	<i>n-</i>	<i>no-bol</i>	<i>ni-wit</i>	<i>n-il</i>	<i>n-obon</i>
1INCL/2	<i>V-</i>	<i>o-bol</i>	<i>i-wit</i>	$\emptyset$ - <i>il</i>	$\emptyset$ - <i>obon</i>
3AN	<i>g-</i>	<i>go-bol</i>	<i>gi-wit</i>	<i>g-il</i>	<i>g-obon</i>
REFL	<i>d-</i>	<i>do-bol</i>	<i>di-wit</i>	<i>d-il</i>	<i>d-obon</i>
RECP	<i>t-</i>	<i>to-bol</i>	<i>ti-wit</i>	<i>t-il</i>	<i>t-obon</i>

The Bunaq agreement system has a host of lexical complexities. There are seven different conjugational classes of verbs, chiefly based on inflectional behaviour in the third person. The class with the largest membership is the zero conjugation; this has no agreement prefix coindexing a third person inanimate P and the prefix *g-* for a third person animate P. The remaining conjugations are distinguished by the initial consonants which are present on the verb with a third person inanimate P. As illustrated in Table 14, these initial consonants are replaced by the *g-* prefix when the P is animate. Not all verbs that have these consonants show this replacement. Conjugation classes are therefore not the result of predictable morphophonemic processes, but rather their membership is a lexical property of verbs. Most of the initial consonant conjugation classes contains many dozens of members, involving therefore a considerable amount of learning on the part of speakers.

Table 14: Conjugation classes of Bunaq verbs

	INANIMATE	ANIMATE	
zero conjugation	<i>iwal</i>	<i>giwal</i>	‘pick’
	<i>teke?</i>	<i>geteke?</i>	‘look at’
h-conjugation	<i>hukat</i>	<i>gukat</i>	‘lift’
s-conjugation	<i>sumi</i>	<i>gumi</i>	‘hide’
t-conjugation	<i>tinik</i>	<i>ginik</i>	‘cook’
d-conjugation	<i>doenik</i>	<i>goenik</i>	‘remember’
l-conjugation	<i>logo</i>	<i>gogo</i>	‘move sth’

Bunaq also has various morphophonemic rules associated with agreement prefixes. Notable amongst these is the metathesis of roots of the shape  $CV_1V_2C$  under prefixation

where  $V_1$  is high and  $V_2$  non-high to the shape  $-V_1'CV_2C$ . Table 15 presents two examples, one noun *luel* and one verb *sie?*.

Table 15: Examples of Bunaq metathesis under prefixation

		<i>luel</i> ‘peel’	<i>sie?</i> ‘tear’
1EXCL	<i>n-</i>	<i>n-ulel</i>	<i>n-ise?</i>
1INCL/2	<i>V-</i>	$\emptyset$ - <i>ulel</i>	$\emptyset$ - <i>ise?</i>
3AN	<i>g-</i>	<i>g-ulel</i>	<i>g-ise?</i>
REFL	<i>d-</i>	<i>d-ulel</i>	<i>d-ise?</i>
RECP	<i>t-</i>	<i>t-ulel</i>	<i>t-ise?</i>

Numerous irregular root mutations are found on Bunaq verbs under prefixation. The changes in roots typically involve the duplication of a segment or a segment’s deletion, as illustrated in Table 16. The changes are not predictable based on the shape of the root.

Table 16: Examples of Bunaq irregular verb root mutation under prefixation

	Unprefixed form	3 <sup>rd</sup> person prefixed form
‘split’	<i>bagal</i>	<i>gagabal</i>
‘gather’	<i>binun</i>	<i>gibibun</i>
‘wash’	<i>ili</i>	<i>gigili</i>
‘tell’	<i>pila?</i>	<i>gipiala</i>
‘steal’	<i>bini</i>	<i>gibi</i>
‘beat’	<i>tu?u</i>	<i>gutut?</i>
‘stretch’	<i>mene</i>	<i>gemen</i>
‘clear’	<i>naman</i>	<i>gaman</i>
‘cover’	<i>bolok</i>	<i>gobok</i>
‘break’	<i>pili</i>	<i>gipi</i>

In Makalero and Makasae, agreement similar to that found in Bunaq is limited to the 3<sup>rd</sup> person. Cognates of the Bunaq 3<sup>rd</sup> person inflection *g-* are found on numerous vowel-initial items. Makalero has a set of verbs and preverbs that inflect for *k-* marking a 3<sup>rd</sup> person P argument (for more details on the morphosyntactic conditions of *k-*, see Huber 2011: 349ff). Makasae has a fossilised *g-* which is a reflex of the same prefix. Table 17 presents the set of cognates between Makalero and Makasae where reflexes of *\*g-* are found.<sup>12</sup>

From Table 17 we see that the Makasae cognates of the Makalero inflecting verbs have, for the most part, fossilised the prefix and thus almost always have initial /g/. There

<sup>12</sup> These are not the only items with a reflex of *\*g-* on them in Makalero and Makasae. There are several other inflecting verbs in Makalero, but they do not have known cognates in Makasae. They are: *ako*, *k-ako* ‘steal’, *asu*, *k-asu* ‘for’, *ati-*, *k-ati-* ‘downwards’, *horu*, *ko-horu* ‘with’, *uan-*, *k-uan-* ‘bigger, more’, *uri-*, *k-uri-* ‘obscured, hidden’, *uta-*, *k-uta-* ‘hidden from view’, *a?a-*, *k-a?a-* ‘onto’. There are also several *g-*-initial items in Makasae which look like they may have a fossilised *g-*. They are: *goba* ~ *guba* ‘with’, *gamun(u)* ‘hold in hand’, *gume* ‘pick fruit’, *guhur(u)* ‘blow’.



are few exceptions to this fossilisation pattern, such as Makasae *amuʔu* ‘smell, stink’. In some cases, semantics seems to play a role in prefix fossilisation. With *gapu* ‘with’ and *apu* ‘carry, cradle’, Makasae retains both prefixed and unprefixed forms, albeit with differing semantics (similar to Makalero). Makasae *umu* ‘die’ has no fossilised prefix, likely because inflected forms had a transitive meaning ‘kill’ (cf. Makalero *k-umu-* ‘to kill (someone by an action’, Bunaq *g-ume* ‘kill’). Such a transitivity function of \*g- is also apparent on the Makasae verb *gira* ‘to water’ < *g-* + *ira* ‘water’ (n).

Table 17: Comparison of appearance of reflexes of \*g- prefix on cognate roots in Makalero and Makasae

Makalero				Makasae			
vowel form		k-form		vowel form		g-form	
<i>ali-</i>	‘all over’	<i>kali-</i>		--		<i>gali</i>	‘back, around’
<i>afa-</i>	‘away from’	<i>kafa-</i>		--		<i>gafa</i>	‘from, leave behind’ <sup>†</sup>
<i>afi-</i>	‘sideways’	<i>kafi-</i>		--		<i>gafi</i>	‘beside, next to, across’
<i>afu</i>	‘carry, with’	<i>kafu</i>	‘carry a child’	<i>apu</i>	‘carry, cradle’	<i>gapu</i>	‘with, for, bring, take’
<i>amu?</i>	‘smell’	<i>kamu?</i>		<i>amu?u</i>	‘smell, stink’	--	
<i>ata-</i>	‘in contact’	<i>kata-</i>		--		<i>gata</i>	‘next to, beside, near’
<i>e-</i>	‘firm’	<i>ke-</i>		--		<i>ge-</i>	‘firm’
<i>ena</i>	‘see’	<i>kena</i>		<i>ena</i>	‘see, look, watch’	<i>(nehe)gena</i>	‘see from afar’
<i>ene</i>	‘hit, strike’	<i>kene</i>		--		<i>gene</i>	‘hit, afflict’
<i>eta-</i>	‘apart’	<i>keta-</i>		--		<i>geta</i>	‘apart’
<i>ha?awein</i>	‘place’ (n)	--		--		<i>ga?awai</i>	‘place’ (n)
<i>ia-</i>	‘under’	<i>kia-</i>		--		<i>gia</i>	‘under, inside’
<i>ini</i>	‘do, make’	<i>kini</i>		--		<i>gini</i>	‘do, make, give’
<i>ira</i>	‘water’	--		<i>ira</i>	‘water’	<i>gira</i>	‘to water’, ‘juice, internal liquid’
<i>isa</i>	‘bake, roast’	<i>kisa</i>		--		<i>gisa</i>	‘roast’
<i>isi-</i>	‘be at’	<i>kisi-</i>	‘originate, belong’	<i>isi</i>	‘at, in, to, from, since’	--	
<i>ou-</i>	‘towards’	<i>kou-</i>		--		<i>gau</i>	‘for, towards, at’
<i>ouar</i>	‘master, owner’	--				<i>gauhaa</i>	‘master, owner’
<i>ua-</i>	‘top’	<i>kua-</i>		--		<i>gua</i>	‘on (top of), over’
<i>ue-</i>	‘around’	<i>kue-</i>		--		<i>goe</i>	‘around’
<i>umu-</i>	‘die, dead’	<i>kumu-</i>	‘to kill by an action’	<i>umu</i>	‘dead, die’	--	
<i>uta</i>	‘kill’	<i>kuta</i>		--		<i>guta</i>	‘kill’
<i>utu</i>	‘cover, block, wear’	<i>kutu</i>		--		<i>gutu</i>	‘put on, wear, bar’
<i>utu?</i>	‘mind, look after’	<i>kutu?</i>		--		<i>gutu</i>	‘to look after, keep eye on’

<sup>†</sup> The only instance of this item in Correia (2011: 74) is glossed as postposition the following: *i bese-bese tagara bo mikorlet i hau gafa ria?a* 2PL quickly walk so minibus 2PL PERF POSP run ‘You must walk quickly, otherwise you will miss the minibus’. The meaning of *gafa* here appears to be malefactive. Huber (p.c.) has one example of *gapa* from the Makasae Ossu dialect where *p* and *f* have not fully merged: *ini wata=e na?u gapa la?a* 1PL.INCL coconut=DEF just from go ‘We left, leaving the coconuts behind’.

In Makasae we can also see a tendency for the *g-* prefix to be retained in contexts where it is ‘trapped’ between morphemes. This is apparent from the Makasae verb *ena* ‘see, look, watch’, which when marked with the intensifying prefix *nehe-* exhibits the prefix *nehe-gena* ‘see from afar’. “Trapped” *g-* is also found fossilised on nouns in Makasae with inalienable semantics in other compounds. For example, Makasae *gaʔawai* ‘place’ occurs in compounds such as *basara-gaʔawai* ‘marketplace’ and *tana-gaʔawai* ‘fingerprint’ (lit. hand place), or Makasae *gauhaa* ‘master, owner’ which occurs in compounds such as *oma-gauhaa* ‘host’ (lit. house owner) and *keta-gauhaa* ‘farmer’ (lit. ricefield owner).<sup>13</sup> Similarly, an inflected form of *ira* ‘water’ is also found in the Makasae compound *awa-gira* ‘penis water’ for ‘sperm’ (cf. Bunaq *g-il* 3-water for ‘juice (of a fruit), bodily fluid’). Makalero does not appear to retain reflexes of \**g-* on nouns, instead using a newly grammaticalised form of the third person pronoun *ki* (cf. Makasae *gi* ‘3’) in its place.

The evidence from Makasae and Makalero indicates that in their immediate common ancestor the third person agreement \**g-* was still used productively on vowel-initial (and some h-initial) nouns and verbs to mark inalienable possessors and P arguments respectively. By contrast, in their closest relatives, Fataluku and Oirata, \**g-* has not been preserved. However, there is evidence that a cognate agreement prefix was present in the common ancestor of these languages.

Fataluku has a prosthetic vowel that coreferences an argument on consonant initial verbs. The vowel is found on a restricted set of verbs and is a lexical property of those verbs. Typically, the prosthetic vowel is a copy of the first vowel of the root. However, in a few cases, the prosthetic vowel is unpredictable. Example verbs, both with and without prosthetic vowel, are set out in Table 18.

Table 18: Fataluku verbs with vowel prothesis (irregular prosthetic vowels bolded)

<b>a</b>	‘send’	<i>har</i>	<i>ahar</i>	<b>o</b>	‘embrace’	<i>kolev</i>	<i>okolev</i>
	‘flare’	<i>kan</i>	<i>akan</i>		‘recognise’	<i>nof</i>	<i>onof</i>
	‘multiply’	<i>ruka</i>	<b><i>aruka</i></b>		‘cut up’	<i>fof</i>	<i>ofof</i>
	‘warp’	<i>ha</i>	<i>aha</i>		‘be.inside.PL’	<i>fo</i>	<i>ofo</i>
<b>e</b>	‘take’	<i>me</i>	<i>eme</i>		‘be.inside.SG’	<i>to</i>	<i>oto</i>
	‘read’	<i>ler</i>	<i>eler</i>	<b>u</b>	‘clothe’	<i>lavere</i>	<b><i>ulavere</i></b>
	‘count’	<i>keh</i>	<i>ekeh</i>		‘sweep’	<i>lur</i>	<i>ulur</i>
	‘wipe’	<i>fer</i>	<i>efer</i>		‘catch in hand’	<i>nam</i>	<b><i>unam</i></b>
	‘measure’	<i>te</i>	<i>ete</i>		‘grab’	<i>fal</i>	<b><i>ufal</i></b>
<b>i</b>	‘cook anew’	<i>tih</i>	<i>itih</i>		‘feed’	<i>fan</i>	<b><i>ufan</i></b>
	‘wait’	<i>hir</i>	<i>ihir</i>		‘plant’	<i>tu</i>	<i>utu</i>
	‘bind’	<i>sil</i>	<i>isil</i>		‘spoon up’	<i>huleve</i>	<i>uhuleve</i>

The morphosyntactic properties of the Fataluku prosthetic vowel are very similar to the properties displayed by agreement prefixes in its relatives. The prosthetic vowel in Fataluku coindexes P on transitive verbs and S on intransitive verbs.<sup>14</sup> This parallels the split-S

<sup>13</sup> These nouns are frequently inalienably possessed in Timor languages, for instance, Bunaq *g-omo* ‘owner, master’ and *g-oloʔ* ‘place, spot’.

<sup>14</sup> Often the prosthetic vowel is anaphoric, referring back to an earlier established argument that is elided in the clause with the prosthetic vowel, but it is by no means restricted to contexts where a referent is not expressed by independent nominal constituents.

distribution of agreement prefixes in the related languages of Alor and Pantar as well as Bunaq. Marking a verb with a prosthetic vowel prevents another prefix from being added to the verb and never triggers initial consonant mutation (see Section 5.4 where this process is discussed as the result of the historical loss of a prefix \*n-). These features indicate that the prosthetic vowel is not simply the result of a morphophonemic process, but rather fills the prefix slot on the verb. In Bunaq and Makalero an agreement prefix also cannot co-occur with locative *n-*. These features indicate that what has been described as a prosthetic vowel in Fataluku is in fact an agreement prefix *V-*.

A prefix of this form would be the expected reflex of the Proto-Timor-Alor-Pantar (PTAP) \*ga- ‘3’ (reflected in Proto-Alor-Pantar \*ga-; Proto-Maka (the common ancestor of Makasae and Makalero) \*g- > Makasae fossil *g-*, Makalero *k-*; Bunaq *g-*. PTAP \*g > \*ʔ medially in Proto-Eastern Timor (PET), but was retained as \*g initially. These phones were continued with the same values in Proto-Maka, but in Proto-Frata (the common ancestor of Fataluku and Oirata), the sound change progressed with PTAP \*g > \*ʔ in all positions. However, glottal stop is not contrastive in initial position in Fataluku, thus leaving only an initial vowel as a possible reflex of PTAP \*ga- in Fataluku. The remaining vowel reflex of \*ga- became unspecified *V-* either in Fataluku or at an earlier stage.<sup>15</sup> A parallel case supporting the change of \*g > \*ʔ > Ø initially in Fataluku is the PET pronoun \*gi ‘3.POSS’, reflected as Makasae *gi*, Makalero *ki* and Fataluku *i*.

The irregular prefixal vowels that we find on some Fataluku verbs appear to be the result of reanalysis of root vowels as prefixes. For example, the irregular prefixal vowel *u-* on Fataluku *nam-e* ‘catch in hand-VBLZ’ goes back to PTAP \*amun ‘grab’, a form which is reflected as Makasae *g-amun* ‘hold in hand’ and Bunaq *amu?* ‘seize’. The PTAP form metathesised \*unam- in Proto-Frata. The initial \*u vowel was reanalysed as an agreement prefix leaving the modern-day root as *nam-* in Fataluku and Oirata. The same reanalysis can also be seen in Fataluku *laver-e* ‘clothe-VBLZ’, a verb derived from the Fataluku noun *ulavari* ‘waist’.<sup>16</sup> The ‘prefixed’ form of the verb, *ulaver-e*, reflects the original initial vowel segment of the root still preserved on the nominal root.

In sum, like the Austronesian languages, the Papuan languages of Timor are not without verbal inflectional morphology. Bunaq has the most extensive paradigm of agreement prefixes. Prefixation in Bunaq is also associated with lexicalised conjugation classes, morphophonemic processes like metathesis, and a host of irregular root changes. While agreement prefixes are reduced in the Eastern Timor Papuan languages, they have by no means been stripped from the languages. Makalero and Makasae preserve reflexes of PTAP \*ga- on vowel initial verbs and nouns. In Fataluku a reflex of the same agreement prefix *V-* appears on a defined set of consonant initial verbs. Reanalysis of root initial vowels as instantiations of the *V-* prefix have also led the prefixal vowel to have an unpredictable form on numerous verbs.

---

<sup>15</sup> Recall that Makalero and Makasae only have retained agreement prefixes on vowel initial verbs and hence the vowel is lost.

<sup>16</sup> Cf. Oirata *ulawara* ‘waist, loins’. The verb *laver-e* ‘clothe-VBLZ’ is likely to have originally meant ‘put on a loincloth’ or ‘wrap cloth around waist’. The initial *ula-* of this item reflects PTAP \*[w]ula ‘tail’ and is found in numerous complex nominals in Fataluku and Oirata, e.g., Fataluku *ulafuka* ‘tail’, *vehula* ‘youngest child’

## 5.2 Animacy and agreement

The Papuan languages of Timor manifest animacy distinctions on a range of agreement targets.

Some lower numerals in Eastern Timor Papuan languages have different agreement forms for human versus non-human referents. In Makalero, this semantic distinction applies to ‘two’ ‘three’ and the quantifier ‘many’ (Table 19). In Makasae, ‘two’, ‘three’ and ‘four’ have distinct forms for human and non-human referents (Table 20). Human numerals generally combine with a human classifier, Makalero *amu* ‘body’ and Makasae *anu* ‘person’. The formation of these distinct numerals does not follow any known inflectional pattern in the languages, but does appear to involve fossilised prefixes on the numeral bases (e.g., \*lol- for non-human, \*mV- for human).

Table 19: Makalero numeral agreement

	NON-HUMAN	HUMAN
‘one’	<i>u(n)</i>	
‘two’	<i>loloji</i>	<i>meih</i>
‘three’	<i>lolitu</i>	<i>itu</i>
‘four’	<i>fat</i>	
‘five’	<i>lima</i>	
‘six’	<i>douh</i>	
‘seven’	<i>fitu</i>	
‘eight’	<i>afo</i>	
‘nine’	<i>siwa</i>	
‘ten’	<i>ru(ru)</i>	
‘many’	<i>roual</i>	<i>rial</i>

Table 20: Makasae numeral agreement

	NON-HUMAN	HUMAN
‘one’	<i>u</i>	
‘two’	<i>lola?e</i>	<i>mahe</i>
‘three’	<i>lolitu</i>	<i>mitu</i>
‘four’	<i>loloha</i>	<i>pae</i>
‘five’	<i>lima</i>	
‘six’	<i>daho</i>	
‘seven’	<i>pitu</i>	
‘eight’	<i>apo</i>	
‘nine’	<i>siwa</i>	
‘ten’	<i>ruru (u)</i>	
‘many’	<i>baun</i>	

Fataluku also exhibits animacy based agreement on the numerals for ‘two’ and ‘three’ (Table 21). These numerals can be marked with *-afu* when reference is to non-humans and *-tere* when reference is to humans (see Fataluku plural markers discussed in Section 5.4). Unlike Makalero and Makasae, Fataluku also has unmarked numerals for these values which can be used for any referent.<sup>17</sup>

<sup>17</sup> The Fataluku system looks like a reduction of a previously more productive system that has broken down. De Josselin de Jong (1937: 195-197) describes the numeral system of the closely related Oirata language. In this work we see that numerals above one have two forms, one with *-een* and one without. He does not identify any animacy difference between the two forms, though it is notable that numerals marked with *-een* only appear with human nouns in the examples provided. He also notes that the form *-apu* is also found with numerals in Oirata, where it means ‘all’, likely related to Fataluku *-afu*.

Table 21: Fataluku numeral agreement

	UNMARKED	NON-HUMAN	HUMAN
‘one’	<i>ukani</i>		
‘two’	<i>etse</i>	<i>etsafu</i>	<i>etsatere</i>
‘three’	<i>utu?e</i>	<i>utu?afu</i>	<i>utu?atere</i>
‘four’	<i>fate</i>		
‘five’	<i>lime</i>		
‘six’	<i>neme</i>		
‘seven’	<i>fitu</i>		
‘eight’	<i>kafa</i>		
‘nine’	<i>siva</i>		
‘ten’	<i>ta?ane</i>		

Bunaq has a gender system based on a two-way class distinction of INANIMATE versus ANIMATE. Gender is a covert property of Bunaq nouns that is reflected on two agreement targets, determiners and 3<sup>rd</sup> person prefixes on verbs. Determiners in Bunaq must agree in animacy with the head noun of the NP, and each determiner has both an INANIMATE and ANIMATE agreement form (Table 22). On verbs, Bunaq displays differential marking of Ps based on animacy. While INANIMATE Ps are unmarked by a verbal prefix, animate Ps are prefixed on the verb with *g-* ‘3AN-’ (see Table 14 in Section 5.1 for the verbal conjugations which complicate this basic agreement system).

Table 22: Animacy agreement on Bunaq determiners

	INANIMATE	ANIMATE
Definite article	<i>ba</i>	<i>bi</i>
Proximal demonstrative	<i>bare</i>	<i>bari</i>
Non-proximal demonstrative	<i>ba?a</i>	<i>ba?i</i>
Specifier demonstrative	<i>doe</i>	<i>doi</i>
Contrastive demonstrative	<i>homo</i>	<i>himo</i>
Counter-expectational demonstrative	<i>bere</i>	<i>beri</i>

The Bunaq animacy agreement system is described here as a gender system because animacy is a grammatical rather than semantic property of nouns in Bunaq. That is, although the agreement system has a strong semantic basis, it is not sufficient to know the meaning of a noun in order to determine what agreement form it will take. Whilst all nouns denoting animates take ANIMATE agreement, not all nouns denoting inanimates take INANIMATE agreement. For example, *zap* ‘dog’ in (19a) is determined by the ANIMATE form of the definite article and takes the 3<sup>rd</sup> person ANIMATE agreement prefix *g-* on the verb *teke?* ‘look at’. The INANIMATE noun *zo* ‘mango’ in (19b) is determined by the ANIMATE form of the definite article and does not agree on the verb. By contrast, the ANIMATE noun *pa?ol* ‘maize’ in (19c) has a plant referent but takes the ANIMATE agreement forms that we saw with *zap* ‘dog’ in (19a).

- Bunaq gender agreement
- (19) Animate referent with ANIMATE agreement
- a. *neto zap bi ge-teke?*  
 1SG dog DEF.AN 3AN-look.at  
 ‘I’m looking at the dog.’
- Inanimate referent with INANIMATE agreement
- b. *neto zo ba teke?*  
 1SG mango DEF.INAN look.at  
 ‘I’m looking at the mango.’
- Inanimate referent with ANIMATE agreement
- c. *neto paʔol bi ge-teke?*  
 1SG maize DEF.AN 3AN-look.at  
 ‘I’m looking at the maize.’

In sum, animacy plays a role in agreement across the Papuan languages of Timor. Contrastive numerals for human versus nonhuman referents appear irregularly in the Eastern Timor languages. The grammatical gender system of Bunaq with its two-way agreement contrast qualifies in McWhorter’s terms as moderately complex.

### 5.3 Locative and applicative prefixes

Unlike in other TAP languages, verb serialisation is quite limited in Eastern Timor Papuan languages. While other TAP languages have serialisation, similar functions in the Eastern Timor languages are fulfilled by what has variously been described as verbal prefixation, verb compounding and incorporation into a preverbal slot.

Verbal prefixes derived from verbs typically have a dedicated truncated form in Eastern Timor Papuan. We can see this variation between full verb and verb prefix particularly well in Makasae, a language where the move from verb serialisation to verb prefixation is not entirely complete. Comparison in Makasae can be expressed in two ways: (i) the verb *litaka* ‘pass, surpass’ introduces the standard of comparison in serialisation with a stative property verb such as *rau* ‘good’ (20a), or; (ii) the verbal prefix *lita-*, obviously related to the verb *litaka*, introduces the standard as the applied object of the stative property verb (20b). By contrast, Makalero has almost no verb serialisation and its comparatives are formed exclusively by prefixation (Schapper & de Vries 2018).

- Makasae
- (20) Serialised exceed comparative
- a. *fi welafu ehani rau fi boba laneʔe gigeʔe litaka*  
 1PL.INCL life now good 1PL.INCL father PL POSS EXCEED  
 ‘Our lives nowadays are better than our parents’ lives.’ (lit. ‘Our lives now are good exceeding those of our parents’)
- Incorporated/prefixed exceed comparative
- b. *fi welafu ehani fi boba laneʔe gigeʔe lita-rau*  
 1PL.INCL life now 1PL.INCL father PL POSS EXCEED-good  
 ‘Our lives nowadays are better than our parents’ lives.’ (Correia 2011: 318)

The move away from serialisation to prefixation has meant that the Eastern Timor Papuan languages have developed large inventories of verb prefixes, unparalleled in their relatives. Table 23 sets out a small number of the locative prefixes that are found on verbs in Fataluku.

See Section 5.4 for the related issue involving the prefixation of locative verbs marked by *-ne* and the initial consonant mutations on verbs described in the following.

Table 23: Examples of Fataluku (Eastern) prefixes and related *-ne* marked verbs

	Prefixal form	Verbal form		Prefixal form	Verbal form
‘on’	<i>a-</i>	<i>ane</i>	‘at’	<i>metse-</i>	<i>metsene</i>
‘reach’	<i>atsa-</i>	<i>atsane</i>	‘on’	<i>mitsa-</i>	<i>mitsane</i>
‘cover’	<i>atsu-</i>	<i>atsune</i>	‘near’	<i>mini-</i>	<i>minine</i>
‘aside’	<i>afa-</i>	<i>afane</i>	‘in front’	<i>mira-</i>	<i>mirane</i>
‘amidst’	<i>apa-</i>	<i>apane</i>	‘inside’	<i>mutsu-</i>	<i>mutsune</i>
‘far’	<i>tso-</i>	<i>tso</i>	‘around’	<i>poro-</i>	<i>porone</i>
‘front’	<i>fanu-</i>	<i>fanune</i>	‘up, upon’	<i>puhu-</i>	<i>puhune</i>
‘up’	<i>hia-</i>	<i>hiane</i>	‘between’	<i>ulu-</i>	<i>ulune</i>
‘with’	<i>horu-</i>	<i>horune</i>	‘backwards’	<i>uta-</i>	<i>utane</i>
‘under’	<i>iti-</i>	<i>itine</i>	‘away’	<i>ura-</i>	<i>urane</i>

Whilst many of the verbal prefixes in Eastern Timor Papuan languages are productive and have semantics transparently related to their source verb, semantic bleaching of verbal prefixes can also be observed. In these situations, lexicalised relationships between verb and verb prefix tend to emerge. For example, the Makasae verbal prefixes *mi-*, *ne-* and *ge-* are considered by Correia (2011) to be derived from the verb *mini* ‘follow’, *nehe* ‘very, be excessive’ and *gehele* ‘firm, tight’. As can be seen in Table 24, the combination of verbs with these verb prefixes frequently yields unexpected semantics.<sup>18</sup>

Table 24: Examples of Makasae semantically bleached verbal prefixes

<b>mi- prefix</b>			
<i>fusa</i>	‘peep’	<i>mifusa</i>	‘spy’
<i>gamunu</i>	‘hold’	<i>migamu</i>	‘feel, grope’
<i>laʔa</i>	‘go, walk’	<i>milaʔa</i>	‘follow’
<i>loʔi</i>	‘wipe, clean’	<i>miloʔi</i>	‘anoint’
<i>maʔene</i>	‘know’	<i>mimaʔene</i>	‘recognise’
<i>saga</i>	‘look for’	<i>misaga</i>	‘search’
<i>suri</i>	‘let go, set free’	<i>misuri</i>	‘follow in numbers’
<i>tamunu</i>	‘mention’	<i>mitamu</i>	‘name after’
<b>ne- prefix</b>			
<i>akasa</i>	‘try, attempt’	<i>neakasa</i>	‘try hard’
<i>daʔiri</i>	‘flatter, praise’	<i>nedaʔiri</i>	‘flatter, praise greatly’
<i>gini</i>	‘do, make’	<i>negini</i>	‘treat really badly’
<i>guta</i>	‘kill, slaughter’	<i>neguta</i>	‘beat badly’
<i>lolo</i>	‘tell, say’	<i>nelolo</i>	‘scold excessively’
<b>ge- prefix</b>			
<i>base</i>	‘hit, strike’	<i>gebase</i>	‘hammer’
<i>diʔara</i>	‘sit’	<i>gediʔara</i>	‘establish oneself’
<i>booro</i>	‘tie up’	<i>gebooro</i>	‘tighten’
<i>gesi</i>	‘close’	<i>gegesi</i>	‘lock, bolt’
<i>koʔolo</i>	‘hug’	<i>gekoʔolo</i>	‘embrace, include’
<i>sifa</i>	‘hold, drive, catch’	<i>gesifa</i>	‘arrest’

<sup>18</sup> Perhaps it would be more accurate to say that it is not always clear from Correia’s (2011) glosses what the semantic difference between prefixed and unprefixed forms are.



<i>siʔili</i>	‘bind, tie’	<i>gesiʔili</i>	‘fasten, secure’
---------------	-------------	-----------------	------------------

#### 5.4 Initial verb root mutations

The Eastern Timor Papuan languages collectively display irregular initial consonant mutations on verbs that are triggered by what, as we saw in the previous section, has been variously described as prefixation or incorporation into a preverbal slot. For example, Oirata has two patterns of initial consonant mutations:  $t > -r$  (e.g., *tipare* ‘run, flee’ > *ura-ripare* ‘go back, run back’), and  $p > -h$ , (e.g., *pai* ‘make, do’ > *ura-hai* ‘open, uncover’, lit. do back). These mutations are not found on all verbs with the appropriate initial segments; rather they are irregular, non-predictable changes that occur to a verb on prefixation of particular morphemes.

The patterns of initial consonant mutation and the set of verbs subject to them varies from language to language. In Makasae, only around a dozen verbs mutate following two patterns  $s > -d$  and  $t > -d$  (Table 25).<sup>19</sup> In Makalero, the number is larger with around 40 verbs showing consonant mutation following two patterns  $h > -s$  and  $t > -d$  (Table 26).

Table 25: Makasae verbs with initial consonant mutations

	Free	Bound			Free	Bound	
$s > -d$	<i>sege</i>	<i>-dege</i>	‘difficult’	$t > -d$	<i>tamu</i>	<i>-damu</i>	‘name’
	<i>seʔele</i>	<i>-deʔele</i>	‘jump’		<i>taru(nu)</i>	<i>-daru</i>	‘put, place, bury’
	<i>seriki</i>	<i>-deriki</i>	‘tie down’		<i>taʔe</i>	<i>-daʔe</i>	‘sleep’
	<i>sesara</i>	<i>-desara</i>	‘fall, throw down’		<i>tapuru</i>	<i>-dapuru</i>	‘cooked, done’
	<i>sipa</i>	<i>-dipa</i>	‘catch’		<i>tia</i>	<i>-dia</i>	‘bite’
	<i>suri</i>	<i>-duri</i>	‘shoot’				
	<i>sisir</i>	<i>-disir</i>	‘sick’				
	<i>supa</i>	<i>-dupa</i>	‘spit’				

Table 26: Makalero verbs with initial consonant mutations

	Free	Bound			Free	Bound	
$h > -s$	<i>haiʔ</i>	<i>-saiʔ</i>	‘finished’	$t > -d$	<i>tafal</i>	<i>-dafal</i>	‘throw away’
	<i>haka</i>	<i>-saka</i>	‘search’		<i>taka</i>	<i>-daka</i>	‘close’
	<i>hat</i>	<i>-sat</i>	‘dry’		<i>tamu</i>	<i>-damu</i>	‘name’
	<i>haʔal</i>	<i>-saʔal</i>	‘fry’		<i>taru</i>	<i>-daru</i>	‘put, place’
	<i>heil</i>	<i>-seil</i>	‘pull’		<i>tekih</i>	<i>-dekih</i>	‘lean towards’
	<i>hein</i>	<i>-sein</i>	‘wait’		<i>teuh</i>	<i>-deuh</i>	‘buy’
	<i>heti</i>	<i>-seti</i>	‘ask’		<i>teri</i>	<i>-deri</i>	‘cut’
	<i>heke</i>	<i>-seke</i>	‘difficult’		<i>teruʔ</i>	<i>-deruʔ</i>	‘shelter’
	<i>helar</i>	<i>-selar</i>	‘big.PL’		<i>tia</i>	<i>-dia</i>	‘sleep’
	<i>heman</i>	<i>-seman</i>	‘take’		<i>tina</i>	<i>-dina</i>	‘cook’
	<i>heʔel</i>	<i>-seʔel</i>	‘jump’		<i>tiʔal</i>	<i>-diʔal</i>	‘kick’
	<i>heʔi</i>	<i>-seʔi</i>	‘cut’		<i>tiʔ</i>	<i>-diʔ</i>	‘pour (liquid)’
	<i>hifaʔ</i>	<i>-sifaʔ</i>	‘catch’		<i>toʔi</i>	<i>-doʔi</i>	‘dig’

<sup>19</sup> Correia (2011: 216) states that in several cases free and bound forms alternate in the same context, suggesting that initial consonant mutations are continuing to break down in Makasae.

	<i>hofe</i>	<i>-sofe</i>	‘know’		<i>tufa</i>	<i>-dufa</i>	‘sweep’
	<i>hor</i>	<i>-sor</i>	‘protect’		<i>tuku</i>	<i>-duku</i>	‘punch’
	<i>houn</i>	<i>-soun</i>	‘plant’		<i>tula</i>	<i>-dula</i>	‘bring, transport’
	<i>huma</i>	<i>-suma</i>	‘angry’		<i>tule</i>	<i>-dule</i>	‘not want’
	<i>huri</i>	<i>-suri</i>	‘shoot, release’		<i>tupi</i>	<i>-dupi</i>	‘pound, thump’
					<i>tuʔil</i>	<i>-duʔil</i>	‘cook in bamboo’

Fataluku has a similarly sized set of irregular mutating verbs to Makalero, but with more patterns of mutation attested. There are three patterns of initial consonant mutation attested over numerous verbs: *f* > *-p*, *t* > *-ʈ*, and *h* > *-ʈ*, as illustrated in Table 27. A fourth pattern, *s* > *-ʈ*, is known from one verb *sil* > *-ʈsil* ‘bind’.

Table 27: Examples of Fataluku (East) initial mutations

	Free	Bound			Free	Bound	
<i>f</i> > <i>-p</i>	<i>fulu</i>	<i>-pulu</i>	‘spit’	<i>t</i> > <i>-ʈ</i>	<i>tipal</i>	<i>-ʈipal</i>	‘drum’
	<i>fal</i>	<i>-pal</i>	‘grab’		<i>tutef</i>	<i>-ʈutef</i>	‘blow’
	<i>fetil</i>	<i>-petil</i>	‘stumble’		<i>te</i>	<i>-ʈe</i>	‘measure’
	<i>fer</i>	<i>-per</i>	‘wipe’		<i>to</i>	<i>-ʈo</i>	‘inside.SG’
	<i>fo</i>	<i>-po</i>	‘inside.PL’		<i>taja</i>	<i>-ʈaja</i>	‘sleep’
	<i>fot</i>	<i>-pot</i>	‘cut up’		<i>teku</i>	<i>-ʈeku</i>	‘stir’
	<i>fai</i>	<i>-pai</i>	‘do’		<i>tih</i>	<i>-ʈih</i>	‘cook again’
<i>h</i> > <i>-ʈ</i>	<i>hina</i>	<i>-ʈina</i>	‘plait’		<i>tomok</i>	<i>-ʈomok</i>	‘be soft’
	<i>hura</i>	<i>-ʈura</i>	‘spoon’		<i>tu</i>	<i>-ʈu</i>	‘feed’
	<i>hit</i>	<i>-ʈir</i>	‘wait’		<i>tul</i>	<i>-ʈul</i>	‘sick, lazy’
	<i>here</i>	<i>-ʈere</i>	‘dry’				
	<i>ha</i>	<i>-ʈa</i>	‘warp’				
	<i>huleve</i>	<i>-ʈuleve</i>	‘plant’				

These consonant mutations go back to a locative morpheme \**n-* that can in some cases still be observed on some vowel initial verbs under the same conditions as the initial consonant mutations. Both Oirata and Fataluku have a sizeable number of (often cognate) verbs that appear with *n-* when prefixed, e.g., Oirata *asi* ‘see, look’ > *ura-nasi* ‘look back’ and *ihile* ‘fly’ > *ura-nihile* ‘fly back’, Fataluku *atsi* ‘see’ > *mucu-natsi* ‘look inside’, *ipile* ‘fly’ > *mucu-nipile* ‘fly inside’. Makalero preserves this *n-* on a small number of verbs also: *umu* ‘die’ > *afa-numu* ‘die leaving (so.) behind’ (cf. *k-umu* 3-die ‘kill so.’), *uta* ‘fall (of rain)’ > *isi-nuta* ‘fall on (of rain)’, *k-ini* ‘make’ > *ata-nini* ‘make in’. The initial consonant mutations seen above are consistent with assimilation to an earlier \**n-* prefix that has subsequently been lost except on (some) vowel initial verbs. The mutations most typically involve a change from a voiceless to a voiced consonant at the same place of articulation, a common result of nasalisation.<sup>20</sup> The changes in place of articulation (e.g., *h* > *-s*) are also consistent with the presence of alveolar nasal.

<sup>20</sup> Note that Fataluku /ʈ/ corresponds to Makalero/Makasae /d/ and appears to represent a sound change from an earlier voiced consonant reconstructed by Schapper et al. (2014) as \*D.

The reconstruction of \**n-* as a locative morpheme is motivated by the fact that the most common trigger for the initial consonant mutations and the *n-* prefix on vowel initial verbs is a prefix with locative semantics. Additional evidence for locative \**n-* comes from Bunaq, where a subset of obligatorily possessed nouns show a contrast between locative *n-* and third person agreement marker *g-*.<sup>21</sup> The locative prefix on these items in Bunaq expresses that the referent of the noun it marks has an internal location. For example, on *-iol* ‘voice, sound, language’, *n-* in (21a) indicates that the noise issues from the engine internal to the motorbike, whereas *g-* in (21b) marks simply that the sound is that of a motorbike, e.g., heard from a distance. Similarly, on *il* ‘water’, *n-* in (21c) indicates the reference is to the water internal to the nut of the coconut palm, while *g-* in (21d) denotes the fluid circulating through the palm, i.e., its sap.

Bunaq contrast of *n-* and *g-* prefixes on nouns

(21)	<i>n-</i> form		<i>g-</i> form
	a. <i>motor</i> <b><i>n-iol</i></b>		b. <i>motor</i> <b><i>g-iol</i></b>
	motorbike LOC-voice		motorbike 3-voice
	‘internal growl of motorbike engine’		‘motorbike’s sound’
	<i>n-</i> form		<i>g-</i> form
	c. <i>hoza</i> <b><i>n-il</i></b>		d. <i>hoza</i> <b><i>g-il</i></b>
	coconut LOC-water		coconut 3-water
	‘water contained within a coconut’		‘sap of a coconut tree’

Here again we have a situation of prefixes that have been accreted and created new complexity in the form of irregular mutations on some initial consonants on verbs and the availability of an *n-* prefix on other roots.

### 5.5 Derivational suffixal morphology

Fataluku and Oirata have the largest array of derivational morphology within the Timor-Alor-Pantar languages. Whilst, as already mentioned, derivational morphology does not hold the same status as inflectional morphology for McWhorter (2008: 18-20), the extent of derivational morphology in some of the Papuan languages of Timor is noteworthy.

Of these, Fataluku has the largest number of documented derivational suffixes, some of which are high-frequency items. Verbs in Fataluku can be nominalised by means of two suffixes: *-n* and *-ana* (or its allomorph *-nana* occurring on vowel-final verbs). The suffix *-n* (often realised as [-nu] at the end of a phonological phrase) only occurs on vowel-final verbs that are not marked with the verbaliser *-e* (examples in Table 28). The nominalising suffix *-ana* can occur on verbs both with and without *-e* (examples in Table 29).<sup>22</sup> These two suffixes

<sup>21</sup> While it may seem far-fetched to relate an *n-* prefix on nouns in one language with one found on verbs in other languages, recall that in Bunaq, as in other Timor-Alor-Pantar languages, the same paradigm of agreement prefixes on both nouns and verbs. Fataluku also has an *n-* prefix on nouns marking a 3<sup>rd</sup> person inalienable possessor. Because the semantic relationship between locative and inalienable is not clear to me, I make no claim about whether this *n-* is related to the locative *n-*.

<sup>22</sup> While van Engelenhoven (2009b) discusses it as a separate morpheme, following Heston (2015) I regard *-ina* as a likely dialect variant of *-ana* chiefly attested in Campagnolo’s (1973) material. Van Engelenhoven (2009b) also describes *-nana* as having adjectival functions. This is because

can occur on the same verbal root with a semantic difference. Van Engelenhoven (2009b) illustrates the contrast between *-n* and *-ana* with the verb *koso* ‘shout’: *koso-n* ‘shout’ refers to the sound produced, while *koso-nana* is an agent nominalisation denoting ‘so. shouting/shouter’.

Table 28: Examples of Fataluku verbs nominalised with *-n*

Verb		Deverbal noun	
<i>atsi</i>	‘see’	<i>atsin</i>	‘vision’
<i>eru</i>	‘lack’	<i>erun</i>	‘shortage’
<i>kolo</i>	‘be mute’	<i>kolon</i>	‘muteness’
<i>toto</i>	‘watch’	<i>toton</i>	‘view’
<i>lika</i>	‘be lean’	<i>likan</i>	‘leanness’
<i>mase</i>	‘eat’	<i>matsen</i>	‘food’
<i>latsa</i>	‘make fence’	<i>latsan</i>	‘fenced place’

Table 29: Examples of Fataluku verbs nominalised with *-(n)ana*

Verb		Deverbal noun	
<i>atsi</i>	‘see’	<i>atsinana</i>	‘sth. seen, visible’
<i>afile</i>	‘slice’	<i>afilana</i>	‘slice, sth. sliced’
<i>akate</i>	‘swollen’	<i>akatana</i>	‘sth. swollen’
<i>ase</i>	‘rough’	<i>asana</i>	‘sth. rough’
<i>tsulu</i>	‘cook’	<i>tsulunana</i>	‘cooked food’
<i>hoile</i>	‘hunt’	<i>hoilana</i>	‘prey, sth. hunted’
<i>itsane</i>	‘fall’	<i>itsanana</i>	‘sth. falling’
<i>isi</i>	‘descend’	<i>isinana</i>	‘sth. descending’
<i>tahine</i>	‘beautiful’	<i>tahinana</i>	‘so. beautiful’

According to van Engelenhoven & Huber (forthcoming) and Heston (2015), nouns ending in a consonant in Fataluku can be verbalised by means of the suffix *-e*. Examples of such precatogorical roots are provided in Table 30.

Table 30: Examples of Fataluku nouns verbalised with *-e* ‘VERB’

Noun		Denominal verb	
<i>lapar</i>	‘chop’	<i>lapare</i>	‘chop’
<i>lamak</i>	‘crumb’	<i>lamake</i>	‘crumble’
<i>iʔis</i>	‘vomit’	<i>iʔise</i>	‘vomit’
<i>tsatan</i>	‘sign’	<i>tsatane</i>	‘sign’
<i>laman</i>	‘orchard’	<i>lamane</i>	‘make an orchard’
<i>apat</i>	‘tuber’	<i>apate</i>	‘emerge from soil’
<i>asir</i>	‘salt’	<i>asire</i>	‘put salt on’
<i>kosin</i>	‘saddle’	<i>kosine</i>	‘saddle (horse)’
<i>hit</i>	‘hit’	<i>hite</i>	‘hit with sword’
<i>tupur</i>	‘woman’	<i>tupure</i>	‘be feminine’
<i>inik</i>	‘sand’	<i>inike</i>	‘be sandy’
<i>matar</i>	‘stone’	<i>matare</i>	‘be stony’
<i>lumuk</i>	‘mud’	<i>lumure</i>	‘be muddy’

---

property denoting verbs marked with *-nana* can also be used in the following nominal attributive function a N<sub>HEAD</sub> i<sub>POSS</sub> V-*nana*<sub>MOD</sub>.

Alongside verbalising *-e*, Fataluku also has the verbalising suffix *-ne* that occurs on vowel final roots. The final *-ne* suffix is dropped when the item is used prefixally on a predicate. Examples of such items have been given already in Table 23. Van Engelenhoven (2009b) analyses *-ne* as an allomorph of *-e* conditioned by the shape of the root. He explains the appearance of *-e* rather than *-ne* on a small number of apparently vowel final roots as due to the presence of an underlying glottal stop (which is, however, lost in some dialects). By contrast, Heston (2015:108-109) argues that the glottal stop that appears on suffixation with *-e* is epenthetic (e.g., *utu* ‘three’ > *utu-e* [utuʔe] ‘three-VBLZ’, *na-* ‘at’ > *na-e* [naʔe] ‘at-VBLZ’) and consequently that the appearance of *-ne* is not morphophonemically predictable. Heston (2015: 24-25) analyses *-ne* as two morphemes *-n* ‘LOC’ and *-e* ‘VBLZ’, adducing that the *-n-e* can be replaced with *-p-e* to give a dynamic (motion) reading (e.g., *mutsu-* ‘inside’, *mutsu-ne* ‘be inside’, *mutsu-pe* ‘go inside’). Whatever the analysis of these suffixes, it is clear that Fataluku makes extensive use of morphological means to derive nouns and verbs (similar morphology is described for Oirata by de Josselin de Jong 1937: 182-183, 188-189).

The other Papuan languages of Timor have far fewer derivational suffixes and they are usually of limited productivity. Makalero, for example, has three derivational suffixes: the first two, *-r* and *-ini*, derive nouns from verbs, illustrated in Tables 31-32.

Table 31: Makalero nominaliser *-ini*

Verb		Noun	
<i>lolo</i>	‘say, speak’	<i>loloini</i>	‘word, conversation’
<i>k-utu</i>	‘wear’	<i>kutuini</i>	‘clothes’
<i>teuh</i>	‘buy’	<i>teuhini</i>	‘gift’
<i>tina</i>	‘cook’	<i>tinaini</i>	‘cooked rice’

Table 32: Makalero nominaliser *-r*

Verb		Noun	
<i>k-ako</i>	‘steal’	<i>akor</i>	‘thief’
<i>nua</i>	‘eat’	<i>nuar</i>	‘food’
<i>umu</i>	‘die’	<i>umur</i>	‘corpse, death’

A further suffix *-ʔ* is used for deriving verbs from nouns, as in the examples in Table 33. This suffix is also found very productively as a verbaliser of verbal prefixes that no longer have an underived verbal form preserved (see Section 5.3).

Table 33: Makalero verbaliser *-ʔ*

Noun		Verb	
<i>uali</i>	‘ear’	<i>ualiʔ</i>	‘hear’
<i>huri</i>	‘brush’	<i>huriʔ</i>	‘brush’
<i>atu</i>	‘faeces’	<i>atuʔ</i>	‘defecate’
<i>teru</i>	‘shelter, umbrella’	<i>teruʔ</i>	‘shelter’

## 5.6 Morphological and suppletive number marking

The Eastern Timor subgroup of Papuan languages is characterised by lexical classes of verbs and nouns that are marked for number by means of a plural suffix or suppletion. Where a suffix is used to mark plural number, its form is often highly irregular.

Verbal number marking in Eastern Timor languages involves different forms being used depending on whether one or more participants are involved in the action. The simplest systems of verbal number marking are found in Makalero and Makasae. Most verbs in these two languages are invariable, but a small class of intransitive verbs has suppletive forms for singular and plural subjects (Table 34 and 35). The plural forms of these verbs typically end in *-ar* ~ *-er* (with a few exceptions in each language), but for the most part there is no clear relationship between singular and plural forms of the roots.

Table 34: Makalero suppletive verbs

	SG	PL
‘sit’	<i>mit</i>	<i>diar</i>
‘stand’	<i>nat</i>	<i>naser</i>
‘run’	<i>ria?</i>	<i>titar</i>
‘lie’	<i>tia, -dia</i>	<i>rou</i>
‘big’	<i>pere</i>	<i>helar</i>

Table 35: Makasae suppletive verbs

	SG	PL
‘sit’	<i>mi</i>	<i>diar</i>
‘stand’	<i>na</i>	<i>nahar</i>
‘run’	<i>ria?</i>	<i>ditar</i>
‘be positioned’	<i>daro</i>	<i>doen</i>
‘lie, rest’	<i>wou</i>	<i>rai</i>
‘say’	<i>lolo</i>	<i>lolini</i>

Fataluku has a much larger set of intransitive verbs marked for singular versus plural. The morphological form the marking takes is unpredictable, as can be seen from the examples in Table 36. The most common form of plural marking is by means of *-re* (on a verb already marked with *-e*) or *-ere*, but many other forms such as *-tere*, *-care*, *-oro* are also found. In addition, a subset of these verbs, mostly posture verbs, has suppletive roots for singular and plural. However, as in Makalero and Makasae, the plural is still typically identifiable by an *-r*.

Table 36: Fataluku verbal number marking and suppletion

	Singular subject	Plural subject		Singular subject	Plural subject
‘fly’	<i>ipile</i>	<i>ipilere</i>	‘dead, die’	<i>umu</i>	<i>umunoro</i>
‘laugh’	<i>kele</i>	<i>kelere</i>	‘live’	<i>lauhe</i>	<i>lauhoru</i>
‘eat’(intr)	<i>matse</i>	<i>matsere</i>	‘slip, be born’	<i>suke</i>	<i>sukoro</i>
‘be located at’	<i>nae</i>	<i>naere</i>	‘stand upright’	<i>tsumai</i>	<i>tsutoro</i>
‘be placed’	<i>hitsine</i>	<i>hitsinere</i>	‘be inside’	<i>otoe</i>	<i>ofoe</i>
‘disappear’	<i>molu</i>	<i>molure</i>	‘hide’ (intr)	<i>palake</i>	<i>pelere</i>
‘full’	<i>polu</i>	<i>polure</i>	‘hang’	<i>vaiake</i>	<i>verire</i>
‘come’	<i>mau</i>	<i>mauere</i>	‘stand’	<i>nate</i>	<i>nehere</i>
‘sleep’	<i>taia</i>	<i>taiatere</i>	‘run, flee’	<i>tifare</i>	<i>helere</i>
‘big’	<i>lafai</i>	<i>lafitsare</i>	‘sit’	<i>mire</i>	<i>tsuare</i>
‘high, long’	<i>lohai</i>	<i>lohitsare</i>	‘lay, lie’	<i>laku</i>	<i>tepere</i>

De Josselin de Jong’s (1937) Oirata materials make it clear that singular ~ plural verbal number marking also exists in that language. Like the other languages of the Eastern Timor group, Oirata has both suppletive (e.g., *mire* / *rua* ‘sit.SG/PL’) and suffixal marking with *-(e)re* (e.g., *naaje* / *naajere* ‘swim.SG/PL’) for the number of the subject of an intransitive verb. In addition, Oirata uses *-(e)re* on transitive verbs to mark the plural number of the subject. Illustration of the number contrast is given in (22).

Oirata plural suffixation of transitive verbs

- (22) Singular subject Plural subject  
 a. *ue in-asi* b. *ite in-asi-ere*  
 3SG 1PL.EXCL-see 2PL 1PL.EXCL-see-PL  
 ‘He sees us.’ ‘You see us.’

Alongside plural number marking on verbs, Eastern Timor languages each have a lexically specified class of nouns denoting humans that take a plural suffix. Plural marking is not obligatory in plural reference in these languages, but it is frequent on this class of human nouns. Examples of members of this special plural marked class are given for Oirata in Table 37. The Oirata plural suffix *-ra* has the allomorph *-a* that appears on nouns which end with a final /r/, as can be seen on the form *tuhur*. There is also an irregular plural form *namirara* derived from *namirai* (cf. Fataluku cognate irregular form in Table 39)

Table 37: Examples of the Oirata restricted plural marked class

	SG	PL
‘younger brother’	<i>noo</i>	<i>noora</i>
‘elder brother’	<i>kaka</i>	<i>kakara</i>
‘sister’	<i>leren</i>	<i>lerenra</i>
‘friend’	<i>hele</i>	<i>helera</i>
‘child’	<i>modo</i>	<i>modora</i>
‘daughter’	<i>modo tuhur</i>	<i>modora tuhura</i>
‘son’	<i>modo nami</i> †	<i>modora namira</i>
‘husband’	<i>namirai</i> †	<i>namirara</i>

† *nami* as an independent noun means ‘man’.

Makalero has a similar lexical class of human nouns marked with a cognate suffix *-raa*. Examples are given in Table 38. This lexically restricted plural suffix exists in Makalero alongside *-laa*, a plural suffix which can mark any noun, though typically it is only used with humans. Makalero also has a dedicated associative plural suffix *-ara* that occurs on personal names.

Table 38: Examples of the Makalero restricted plural marked class

	SG	PL
‘younger sibling’	<i>noko</i>	<i>nokoraa</i>
‘elder sibling’	<i>nana</i>	<i>nanaraa</i>
‘sister-in-law’	<i>ue</i>	<i>ueraa</i>
‘sibling-in-law’	<i>mali</i>	<i>maliraa</i>
‘friend’	<i>pada</i>	<i>padaraa</i>
‘child’	<i>mata</i>	<i>matar ~ mataraa</i> †
‘father’	<i>upa</i>	<i>uparaa</i>
‘uncle’	<i>tiu</i>	<i>tiuraa</i>
‘parent-in-law’	<i>paakin</i>	<i>paakiraa</i>
‘grandparent’	<i>dada</i>	<i>dadaraa</i>
‘widow’	<i>paardufu</i>	<i>paardufuraa</i>

† Huber (2011: 118) gives several permutations of the Makalero plural for ‘children’, including *mata-niki* ‘children’ existing alongside these forms. In Makasae one plural marked noun has been retained but in an irregular form: *mata* ‘child’ and *mata-rini* ‘children’ (Correia 2011) or *mata-kini* (Huber 2008).

Fataluku has a more complex system of plural marking, with multiple small lexical classes of nouns referring to human nouns. The classes are defined by three different plural suffixes: *-r*, *-ra* and *-afu*. Table 39 presents examples of plural suffix taking nouns. We see here that there are also several irregular forms such as those with *-(r)ara* for ‘man’ and ‘woman’. These restricted plural suffixes exist alongside an enclitic =*ere* that can occur on any noun, including on a noun already marked with a plural suffix.

Table 39: Examples of Fataluku restricted plural classes of nouns

	SG	PL
‘father’s sister’	<i>tamu</i>	<i>tamur</i>
‘younger sibling’	<i>noko</i>	<i>nokor</i>
‘older sibling’	<i>kaka</i>	<i>kakar</i>
‘child’	<i>moto</i>	<i>motsor</i>
‘child’	<i>moko</i>	<i>mokor</i>
‘master, owner’	<i>otsava</i>	<i>otsavar</i>
‘friend’	<i>lanu</i>	<i>lanura</i>
‘parent-in-law’	<i>painu</i>	<i>painura</i>
‘sister’	<i>lerenu</i>	<i>lerenura</i>
‘husband’	<i>elehu</i>	<i>elehura</i>
‘wife’	<i>jeu</i>	<i>jeura</i>
‘man’	<i>nami</i>	<i>namira ~ namirara</i>
‘woman’	<i>tupur</i>	<i>tupurara</i>
‘ancestor’	<i>tsal</i>	<i>tsalafu</i>
‘mother’	<i>nal</i>	<i>nalafu</i>
‘father’	<i>pal</i>	<i>palafu</i>
‘person’	<i>mar</i>	<i>marafu</i>

Again, the Papuan languages of Timor have complexity in the domain of plural marking. We observed that McWhorter’s complex features of suppletion, lexicalisation and irregularity all play a role in the expression of plurality across nouns and verbs in these languages.

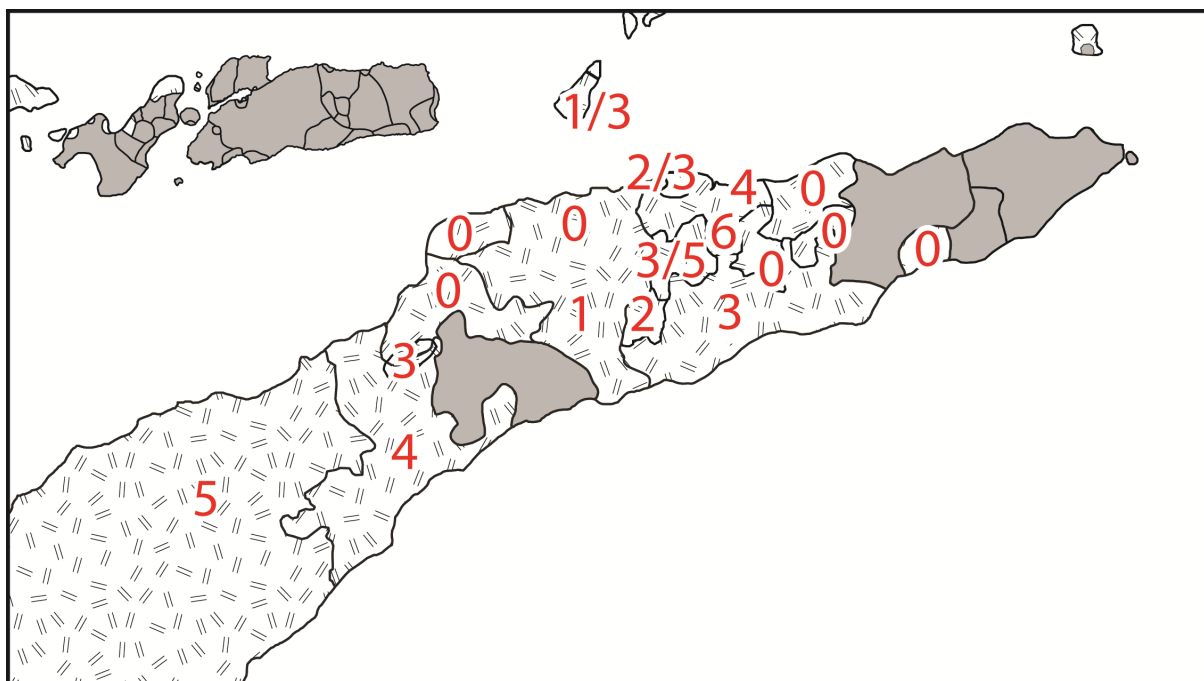
## 6. The comparative picture of complexity in the languages of Timor and surrounds

In the preceding sections, the languages of eastern Timor have been shown to exhibit a range of complexity in the form of morphology, both productive and fossilised. In arguing for his view of complexity, McWhorter stresses that the comparative benchmark for interpreting the degree of complexity displayed by a language is other languages of the same family. He hypothesises that “no language without significant non-native acquisition in its history will be especially simplified in comparison with its sisters” and, relatedly, “all languages with significant non-native acquisition in their histories will be significantly less complex according to my metric than their sisters” (2007: 268). In McWhorter’s view, variation in complexity will be observed amongst sister languages only within a minor range. In this section, therefore, we will take a comparative look at what the languages of eastern Timor look like in contrast to their relatives in the area.



When compared to their relatives in the region, the Austronesian languages of eastern Timor are morphologically somewhat reduced, but not radically so. Reductions are typically limited to small sets of languages and particular types of morphology. For example, eastern Timor Austronesian languages for the most part have very similar sets of subject prefixes to the “complex” ones that McWhorter (2007: 244-246) points to in languages of western Timor, Rotinese and Uab Meto. The complete loss of inflectional prefixes is found only in the neighbouring languages of Tokodede and Kemak as well as in the Kawaimina languages. Between these two groups of languages, we see an assortment of subject prefixes that are not strikingly smaller in distinct number of prefixes than West Timor (Map 2). The loss of subject prefixes here has not been abrupt as McWhorter claims. Simplification of prefixes from \*CV- to simply \*C- appears to have resulted in initial consonant clusters on consonant-initial verbs. A dispreference for these has in turn led to the progressive reduction of subject prefixes. This is seen most pointedly in the fact that subject prefixes tend to be more intact on vowel-initial verbs in eastern Timor languages. At the same time, we see that the innovative agreement prefixes on numerals that have retained their CV-shape are not lost in even the highly isolating Kawaimina languages.

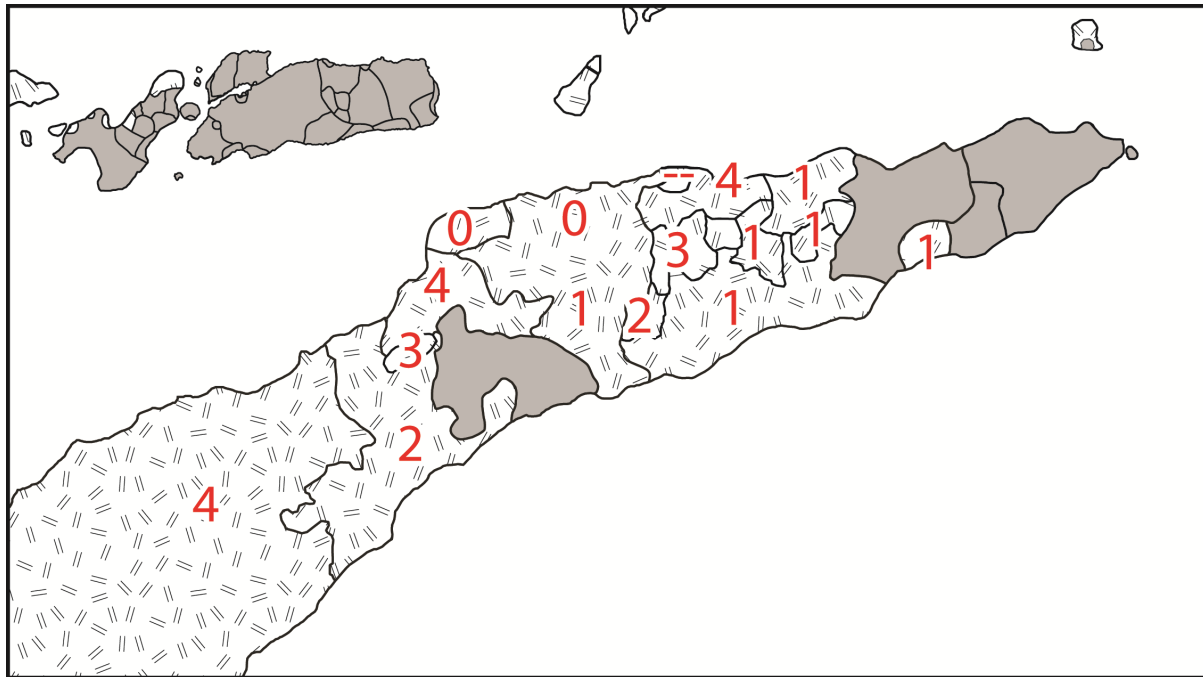
Map 2. Number of verbal subject prefixes per paradigm in Austronesian languages of Timor



Possessive suffixes are in some Austronesian languages of eastern Timor noticeably absent, but it is not the case that these languages have fewer suffixes in general. We see in Map 3 that some eastern Timor Austronesian languages have the same number of possessive affixes as Uab Meto in the west. It is accurate to say that the typical eastern Austronesian pattern of suffixes marking inalienable possessors and free elements marking inalienable possessors (see Donohue and Schapper 2008) has broken down in some languages with the result that fewer affixes are involved in the coding of possessors. Systems of possessive classification have been reformed in different ways with possessive suffixes and free possessor markers entering into paradigmatic relationships in which factors such as phonological shape of the root and person govern possessive marking, not simply

inalienability. At the same time, a language like Kemak preserves the original alienability split and the associated possessive suffixes, while Dadu'a has developed new prefixal encoding of possessors. This diverse coding of possessors in the eastern Timor Austronesian languages points to a cycle of morphological loss and creation that McWhorter sees as part of normal language change.

Map 3. Number of possessive suffixes per paradigm in Austronesian languages of Timor



Synchronic metathesis is a notable morphosyntactic feature of the Austronesian languages Uab Meto spoken in West Timor and Leti spoken on an island off the eastern tip of Timor (see Schapper 2015: 137-138 on its areal distribution). In both these languages, metathesis is conditioned by an intricate set of pragmatic and grammatical contexts, and affects a large portion of the lexicon (see Steinhauer 1993, 1996; Edwards 2016; van Engelenhoven 2004). In McWhorter's (2007: 245) terms, metathesis is a complex feature that requires significant learning of morphosyntactic and morphophonological rules. Whilst synchronic metathesis is not as pervasive in the central Timor languages Mambae and Kemak as it is in Uab Meto and Leti, the fact that it is found with a large number of basic vocabulary items contradicts the idea that these languages are simple. The existence of a typological rarity such as synchronic metathesis in their grammars rather suggests that the significant erosion of inflectional morphology in the central Timorese languages is not the result of any unnatural interruption to language transmission.

McWhorter (2007: 244) admits that the derivational prefixes of Tetun are similar to those found in Rotinese. We saw in Section 4.2 that, whilst not all Austronesian languages in eastern Timor have the same amount of derivational morphology as Tetun, derivational prefixes are not unusual in the languages, in either productive or fossilised form. What is more, in those languages where derivational prefixes are not productive, they have been accreted and, in several cases, have given rise to new phonemes. The most striking example of this is the Kawaimina languages where the pressure to rid the languages of consonant clusters have caused derivational prefixes to be absorbed to create new consonant phonemes,

making one of the largest consonant inventories of the Austronesian family. Crucially, even when fossilised, these prefixes have not been ‘shed’ from the languages, but have created new complexity in other domains of linguistic structure.

Turning to the Papuan languages, we also do not find a significant difference in complexity between the Papuan languages of Timor and their sisters of the Alor-Pantar subgroup. PTAP only had a singular paradigm of inflectional prefixes for person/number, reflexive and reciprocal occurring on both nouns and verbs (Schapper, Huber and van Engelenhoven 2012). This system is still found in the languages of Pantar, but has been expanded in the languages of Alor through the morphologisation of originally free pronominal forms with the shape CV as additional agreement paradigms. Bunaq continues the PTAP system of a single paradigm of agreement prefixes on nouns and verbs, though due to vowel harmony, the number of distinctions marked by prefixal vowels has been lost. In addition, we saw in Section 5.2 that verb prefixation in Bunaq is associated with numerous irregular changes of verb forms and with arbitrary conjugation classes. The agreement system of the Eastern Timor subgroup of Papuan languages shows reduction, with only the PTAP third person prefix having reflexes. The erratic appearance of reflexes, for example, on vowel-initial nouns and verbs in Makasae and Makalero or as an initial vowel of unpredictable shape in Fataluku represents complexity in the form of irregularities that are not found in Alor-Pantar languages.

Whilst Alor-Pantar languages do have more inflectional prefixes than are found in the Papuan languages of Timor, they completely lack the verbal number suppletion or suffixal inflection for plurality found in the Papuan languages of the Eastern Timor group. Although unique to this TAP subgroup, there is good reason not to consider the feature to be new. Papuan languages on New Guinea, where the ancestor of TAP languages presumably originated, frequently have small classes of kin terms that are marked for plural number and of verbs that are suppletive for number, just as in the Eastern Timor Papuan languages. We also find these patterns in the West Bomberai languages (Iha, Mbaham and Kalamang), the relatives of TAP languages on New Guinea. For example, Mbaham has at least two nouns that are marked for number (*namiha* ‘man.SG’ / *namiata* ‘man.PL’, *tumbuhar* ‘woman.SG’ / *tumbuota* ‘woman.PL’), while Kalamang has a suffix *-mur* only found on kin terms (Eline Visser p.c.). Verbs suppletive for number have also been attested: *preh-* ‘speak.SG’ and *ngmbeh-* ‘speak.PL’ and *tomot-* ‘sleep.SG’, *tidlh-* ‘sleep.PL’ in Iha (Donohue 2015), and *wes* ‘go.SG’, *wuru* ‘go.PL’, and *mehena* ‘sit.SG’, *ndigi* ‘sit.PL’ in Mbaham (Flassy, Ruhukael and Rumbrawer 1984). Whilst none of these are cognate with suppletive verbs in ET languages, we do have what appears to be a cognate of the PET verbal plural marker \*-r ‘PL’ in Kalamang. Here we find optional plural marking by means of a suffix *-r* on two verbs, *melelu* ‘sit’, *melelur* ‘sit.PL’, and *na* ‘eat/drink’, *nar* ‘eat/drink.PL’ (Eline Visser p.c.). The fact that these morphological patterns in the Eastern Timor subgroup are consistent with patterns of the New Guinea mainland strongly points to their being a retention of an old pattern that is lost in other members of the family.

Some other morphological complexities present in the Papuan languages of Timor are absent in the Alor-Pantar languages and many of them appear to be innovations. The animacy-based gender system of Bunaq is unparalleled among TAP languages (see Schapper 2010b for a description of the wider pattern). The locative and applicative prefixation of the Eastern Timor Papuan languages, particularly in conjunction with the initial root mutations

triggered by it, is more elaborate and complex than in any other family members. Alor-Pantar languages often have between one and three locative applicative prefixes, but these are comparatively limited in number and frequency (Willemsen 2015). Derivational suffixes are also more extensive in Fataluku and Oirata than in the other TAP languages. In Alor-Pantar languages, there is some derivational morphology, most notably *-naw* in Blagar (Steinhauer 2014: 161-163), but it is not widespread.

On proper consideration of a full data set from the languages of eastern Timor, we find that McWhorter's complexity features of overspecification, structural elaboration and irregularity are well-represented in the morphological properties of the languages. While the eastern Timor languages are often, but by no means always, morphologically simpler than their nearest relatives, they are by no means without morphology. Indeed, far from 'shedding' or 'dropping', in McWhorter's terms, all their morphology, these languages have significant amounts of accreted morphology that, through lexicalisation and fossilisation, have given rise to complexities. What is more, the morphological reductions which have undoubtedly occurred in both the Austronesian and Papuan languages of eastern Timor are not the result of drastic stripping, but are explainable with normal processes of progressive phonological and morphosyntactic change. Finally, morphology and its various manifestations are just one kind of complexity according to McWhorter's metric. It is beyond the scope of this paper to discuss other kinds of complexity, but it is worth pointing out that other complexities are by no means absent in eastern Timor. For example, having more than a two-way demonstrative contrast is complex for McWhorter. In this respect, even the most isolating Timorese languages can be regarded as displaying significant complexity: Naueti has three-way distance marked demonstratives plus a fourth non-familiar demonstrative (Veloso 2016: 47-48); Makalero has a five-way demonstrative distinction involving distance and elevation (Huber 2011: 232-233). In McWhorter's model such features, alongside the morphological complexities discussed at length in the preceding sections, must be taken to reflect the 'oldness' of the languages.

## 7. Discussion

In concluding his treatment of the languages of eastern Timor, McWhorter (2007) explicitly deals with potential alternative hypotheses for the isolating word structure that he encounters there. He writes:

Of course, we might attempt to preserve the chance account in the Timorese [...] languages by supposing that the Papuan-related languages developed in this direction because of contact with the reduced Timorese languages. But this would ascribe a highly unusual degree of structural loss to the contact in question. To be sure, Sprachbund effects can occasion loss modeled on one or more of the languages in contact, [...], or the loss of concordial morphemes [...]. But in all of these cases, a great deal of elaboration was retained: the Balkan languages remain highly inflected, as does Gurnu Baagandji, and the languages of Java retain their derivational morphology and certain inflections. But in the case of the Papuan languages of Timor, we must explain why these four languages have not just lost some inflections of a particular kind, but why they lost so much morphology overall that they stand as strangely analytic, or analytic leaning... (McWhorter 2007: 248)

As we have seen throughout this paper, the languages of eastern Timor, be they Austronesian or Papuan, are not unnaturally reduced. That is, affixes have not simply 'shed' from these languages, but have been retained productively and in various reduced and lexicalised forms.

I have put it that, contrary to McWhorter's (2007, 2008) claims, the significant amount of accreted morphological complexity that can be observed provides ample evidence of the 'oldness' of the languages of eastern Timor. McWhorter's claims can only be made because of a massively facile treatment of the data. McWhorter betrays this perhaps most pointedly of all in the following passage, in which he presents the one and only data point from a Timorese Papuan language in all his writings on Timor (McWhorter 2008: 179): "In contrast to typically inflectional Papuan languages is a Fataluku sentence such as *Ana merkadu mara* (I market go) 'I'm going to the market' [...]. Behold a Papuan language with the typology of Chinese." While the rhetoric here is appealing, the typology of a language cannot be captured in any accurate way with a single clause. It hardly bears saying that careful examination of the facts of whole language systems is required to sustain arguments about morphological profiles and their historical origins.<sup>23</sup>

Nonetheless, I do not dispute that the languages of eastern Timor do lean towards having an isolating typology overall. What is the reason for this shared characteristic? In eastern Timor languages, morphological reduction has almost certainly gone hand in hand with phonological changes that we have seen, namely, vowel harmony in unstressed affixal vowels, loss of unstressed vowels in prefixes and in turn the reduction of consonant clusters.<sup>24</sup> The fact that these processes can be observed across Papuan and Austronesian languages in eastern Timor points to just the mechanism that McWhorter himself notes in the above quote. That is, phonological Sprachbund effects between the languages of eastern Timor have occasioned morphological loss. The idea of a Sprachbund centred on eastern Timor is not new. Numerous features such as the absence of a velar nasal phonemes, flexible genitives (i.e., both GEN N and N GEN word orders), nouns for 'face' and 'name' being used as quantifiers of kinds have been noted, among others (Schapper 2011, 2015; Hull 2001). The move towards isolating word structure is adequately provided for by convergence of the concentric circles of linguistic isoglosses around Timor. In short, not only is the supporting data not to be found in the languages themselves, but there is no need for recourse to a hypothesis involving radical simplification due to 'heavy non-native acquisition' in relatively recent times to explain the existence of isolating structure in eastern Timor.

Yet, language shift or significant non-native acquisition must have been a factor in the Timor region. Both the Austronesian and Papuan languages of the area are the result of migrations, the former out of Taiwan (Bellwood 1997) and the latter out of New Guinea (Hull

---

<sup>23</sup> One reviewer suggests that given the materials available at the time of writing, McWhorter's assumption that the Papuan languages of Timor were isolating was not unreasonable. I do not disagree with describing the overall typology of these Papuan languages as isolating(-leaning). However, that is quite different from the position taken by McWhorter in saying that the languages are morphologically stripped and lacking any features typical of 'old' languages. What is more, some of the morphological patterns which I describe here were indeed present in some of the early sketches that would have been available to McWhorter, e.g., Fataluku plural inflection on kin terms and derivational suffixes are described in the early sketch of Hull (2005). A lack of engagement with the existing sources and a failure to exercise caution in making, particularly historical, arguments about poorly described languages is indefensible.

<sup>24</sup> A further idea to be investigated is that prosodic patterns also contributed to morphological loss. Himmelmann (2010) describes Waima'a as not making phonological use of pitch changes (i.e. lacking lexical tone distinctions as well as postlexical pitch accents), but rather as having phrasal accent on the penultimate syllable. It may be that this prosodical type contributes to phonological erosion at the end of the word.

2004; Ross 2005; Schapper 2017; Usher & Schapper ms). The relative timing of the arrival of these two different groups in the Timor area is not clear, but the evidence, such as it is, suggests that they were not separated by a large period of time, both events occurring somewhere in the mid to late Holocene (see Schapper 2015: 141-142 for a summary). But settlement in the region is much older; Timor was populated by modern humans since at least 40,000BP (Hawkins et al. 2017). Genetic studies of human populations in eastern Indonesia show that the dispersal of the Austronesian and TAP language groups was not associated with thorough-going displacement of pre-existing populations (e.g., Richards, Oppenheimer and Sykes 1998; Mona et al. 2007, 2009; Tumonggor et al. 2014; Gomes et al. 2017). As such, the languages of the earliest peoples are likely to have had an impact on those of the incoming groups.

The deeper linguistic prehistory of Timor is, therefore, interesting for understanding the formation of the morphological typology of the languages today. If we compare the morphological properties of Proto-Timor-Alor-Pantar (PTAP), the common ancestor of the Papuan TAP languages, to its relatives on New Guinea, some notable differences in morphology are apparent. In PTAP, case-marking suffixes on nouns and portmanteau person/TAM marking suffixes on verbs are entirely absent, while they are regular features characterising both PTAP's near relatives on West Bomberai and Trans-New Guinea languages in general. If we assume such suffixes can be reconstructed to a higher level (cf. the suffixal plural discussion in Section 6), their loss in PTAP or its predecessor shows a drastic move away from a suffixing inflectional type. At the same time, PTAP retains reflexes of the TNG verbal agreement prefixes, but whereas PTNG verbal agreement prefixes marked P (Suter 2012), PTAP verbal agreement prefixes appear to have marked not only P but also non-active S. These differences between PTAP and its relatives are consistent with PTAP intruding into an area where, as established in the literature, weakly prefixing-to-little inflectional morphology (Gil 2015) and split-S alignment (Donohue 2004) was the norm. Thus, this observation of the profile change of PTAP would add to the wealth of evidence presented in this volume for the ancient origins of isolating word structure across Indonesia.

## References

- Alcantara, Maressa Xavier. 2015. Descrição fonética e fonológica da língua Idaté do Timor Leste. MA dissertation, Universidade de São Paulo.
- Bellwood, Peter. 1997. *Prehistory of the Indo-Malaysian Archipelago*. Honolulu: University of Hawai'i Press.
- Bickel, Balthasar & Johanna Nichols. 2013. Obligatory Possessive Inflection. In *The World Atlas of Language Structures Online*, Matthew S. Dryer & Martin Haspelmath(eds) Leipzig: Max Planck Institute for Evolutionary Anthropology. URL: <http://wals.info/chapter/58>
- Blust, Robert. 2008. Is There a Bima-Sumba subgroup? *Oceanic Linguistics* 47(1): 45–113.
- Blust, Robert. 2014. Austronesian. In *The Oxford Handbook of Derivational Morphology*, Rochelle Lieber & Pavol Štekauer (eds), 545–557. Oxford: Oxford University Press. URL:

<https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199641642.001.0001/oxfordhb-9780199641642-e-030>.

- Boarccaech, Alessandro. 2013. *Dicionário Hresuk-Português*. URL: [https://issuu.com/kirstygunmao/docs/dicionario\\_hresuk\\_-\\_portugues](https://issuu.com/kirstygunmao/docs/dicionario_hresuk_-_portugues)
- Bowden, John, John Hajek & Nikolaus Himmelmann. nd. A sketch grammar of Waima'a. Manuscript.
- Bowden, John, 2010. Metathesis in Helong. In Yassir Nasunius (ed), *KOLITA* 8: 59–63. Jakarta: Pusat Kajian Bahasa dan Budaya, Unika Atma Jaya.
- Correia, A. J. G. 2011. Describing Makasae: A Trans-New Guinea Language of East Timor. PhD dissertation, University of Western Sydney.
- Campagnolo, Henri. 1973. La langue des Fataluku de Lórehe (Timor Portugais). PhD dissertation, Université de Paris V - René Descartes.
- da Silva, Eng. Guilherme. 2012. *Disionáriu Wekais-Tetun*. Dili: Secretaria de Estado da Cultura.
- Donohue, Mark. 2004. Typology and linguistic areas. *Oceanic Linguistics* 43: 221–239.
- Donohue, Mark. 2015. Verbal Inflection in Iha: A multiplicity of alignments. In *The Oxford Handbook of Inflection*, Matthew Baerman (ed), 405–418. Oxford: Oxford University Press.
- Donohue, Mark & Antoinette Schapper. 2008. Whence the Oceanic indirect possession construction? *Oceanic Linguistics* 47.2: 316–328.
- Edwards, Owen. 2016. Metathesis and Unmetathesis: Parallelism and Complementarity in Amarasi, Timor. PhD dissertation, Australian National University.
- Edwards Owen. 2018a. Top-Down Historical Phonology of Rote-Meto. *Journal of the Southeast Asian Linguistics Society* 11.1: 63–90.
- Edwards Owen. 2018b. Parallel Histories in Rote-Meto. *Oceanic Linguistics* 57(2): 359–409.
- Edwards Owen. 2019. Reintroducing Welaun. *Oceanic Linguistics* 58(1): 31–58.
- Engelenhoven, Aone van. 1987. Preliminaire aantekeningen voor een rekonstructie van het proto-Luangisch-Kisarsch. MA dissertation, Leiden University.
- Engelenhoven, Aone van. 2004. *Leti: A Language of Southwest Maluku*. Leiden: KITLV.
- Engelenhoven, Aone van. 2009a. The position of Makuva among the Austronesian languages in East Timor and Southwest Maluku. In *Austronesian historical linguistics and culture history: a festschrift for Robert Blust*, Adelaar, K. Alexander and Pawley, Andrew (eds), 425–442. Canberra: Pacific Linguistics.
- Engelenhoven, Aone van. 2009b. On derivational processes in Fataluku, a non-Austronesian language in East-Timor. In *The Linguistics of Endangered Languages: Contributions to Morphology and Morpho-Syntax*, 333–362 Leo Wetzels (ed.) Leiden: LOT Occasional Series 13.
- Engelenhoven, Aone van. 2010. Tentatively locating West-Damar among the languages of Southwest Maluku. In *Studia Anthropologica: A Festschrift in Honor of Michael Chlenov*, Svetlana Chlenova and Artem Fedorchuk (eds), 297–326. Moscow-Jerusalem: Gescharim.

- Engelenhoven, Aone van & Juliette Huber. forthcoming. Fataluku. In *Papuan languages of Timor, Alor and Pantar*, Antoinette Schapper (ed), Volume 3. Berlin: Mouton de Gruyter.
- Flassy, Don A.L., Constantinopel Ruhukael & Frans Rumbrawer. 1984. *Struktur Mbahaam Maaq / bahasa Bahaam: Morfologi dan Sintaksis*. Jakarta: Departmen Pendidikan dan Kebudayaan.
- Gil, David. 2015. The Mekong-Mamberamo linguistic area. In *Languages of Mainland Southeast Asia, The State of the Art*, Nicholas J. Enfield & Bernard Comrie (eds), 266–355. Berlin: de Gruyter Mouton.
- Gomes, Sibylle M, Mannis van Oven, Luis Souto, Helena Morreira, Silke Brauer, Martin Bodner, Bettina Zimmermann, Gabriela Huber, Christina Strobl, Alexander W Röck, Francisco Côrte-Real, Walther Parson & Manfred Kayser. 2017. Lack of gene–language correlation due to reciprocal female but directional male admixture in Austronesians and non-Austronesians of East Timor. *European Journal of Human Genetics* 25(2): 246–252.
- Grimes, Charles E., Carlos Marçal & Paolino Fereira. 2014. *Introductory dictionary of Mambae (Same): Mambae—English, English—Mambae, Mambae—Indonesia—Tetun Dili, Indonesia—Mambae, Tetun Dili—Mambae*. Darwin: Australian Society for Indigenous Languages.
- Hawkins, Stuart, Sue O'Connor, Tim Ryan Maloney, Mirani Litster, Shimona Kealy, Jack N.Fenner, Ken Aplin, Clara Boulanger, Sally Brockwell, Richard Willan, Elena Piotto, Julien Louys. 2017. Oldest human occupation of Wallacea at Laili Cave, Timor-Leste, shows broad-spectrum foraging responses to late Pleistocene environments. *Quaternary Science Reviews* 171(1): 58–72.
- Heston, Tyler M. 2015. The Segmental and Suprasegmental Phonology of Fataluku. PhD dissertation, University of Hawai'i.
- Himmelmann, Nikolaus P. 2010. Notes on Waima'a intonation. In *Typological and Areal Analyses: Contributions from East Nusantara*, Michael Ewing & Marian Klamer (eds), 47–69. Canberra: Pacific Linguistics.
- Huber, Juliette. 2008. *First Steps Towards a Grammar of Makasae: A language of East Timor*. (Languages of the World/Materials, 195.) München: Lincom.
- Huber, Juliette. 2011. A Grammar of Makalero. PhD dissertation, Leiden University.
- Hull, Geoffrey. 1998. The basic lexical affinities of Timor's Austronesian languages: a preliminary investigation. *Studies in the Languages and Cultures of East Timor* 1: 97–174.
- Hull, Geoffrey. 2001. A Morphological Overview of the Timoric Sprachbund. *Studies in Languages and Cultures of East Timor* 4: 98–205.
- Hull, Geoffrey. 2002. *Waimaha*. Dili: Instituto Nacional de Linguística.
- Hull, Geoffrey. 2003. *Galoli*. Dili: Instituto Nacional de Linguística.
- Hull, Geoffrey. 2004. The Papuan languages of Timor. *Studies in Languages and Cultures of East Timor* 6: 23–99.



- Hull, Geoffrey. 2005. *Fataluku*. Dili: Instituto Nacional de Linguística.
- Jonker, J. C. G. 1915. *Rottineesche Spraakkunst*. Leiden: Brill.
- Josselin de Jong, Jan Petrus Benjamin de. 1937. *Studies in Indonesian Culture I: Oirata, a Timorese settlement on Kisar*. Amsterdam: Noord-Hollandsche Uitgevers-Maatschappij.
- Klinken, van, Catharina. 1999. *A grammar of the Fehan dialect of Tetun: An Austronesian language of West Timor*. Canberra: Pacific Linguistics.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5.2: 125–166.
- McWhorter, John. 2005. *Defining Creole*. Oxford: Oxford University Press.
- McWhorter, John. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press.
- McWhorter, John. 2008. Why Does a Language Undress? Strange Cases in Indonesia. In *Language Complexity: Typology, Contact, Change*. F. Karlsson, M. Miestamo & K. Sinnemäki (eds), 167–190. Amsterdam: John Benjamins.
- Mills, Roger F. 1991. Tanimbar-Kei: An Eastern Indonesian subgroup. In *Currents in Pacific Linguistics: Papers on Austronesian Languages and ethnolinguistics in Honour of George W. Grace*, Robert Blust (ed), 241–263. Canberra: Pacific Linguistics.
- Mona, Stefano, Mila Tommaseo-Ponzetta, Silke Brauer, Herawati Sudoyo, Sangkot Marzuki, & Manfred Kayser. 2007. Patterns of Y-Chromosome Diversity Intersect with the Trans-New Guinea Hypothesis. *Molecular Biology and Evolution* 24/11: 2546–2555. doi:10.1093/molbev/msm187
- Mona, Stefano, Katharina E. Grunz, Silke Brauer, Brigitte Pakendorf, Loredana Castri, Herawati Sudoyo, Sangkot Marzuki, Robert H. Barnes, Jörg Schmidtke, Mark Stoneking & Manfred Kayser. 2009. Genetic admixture history of eastern Indonesia as revealed by Y-chromosome and mitochondrial DNA analysis. *Molecular Biology and Evolution* 26/8: 1865–1877.
- Penn, David. 2006. Introducing Dadu'a. Honours dissertation, University of New England, Australia.
- Richards, M., S. Oppenheimer & B. Sykes. 1998. mtDNA suggests Polynesian origins in eastern Indonesia. *American Journal of Human Genetics* 63: 1234–1236.
- Ross, Malcolm. 1998. Possessive-like attribute constructions in the Oceanic languages of northwest Melanesia. *Oceanic Linguistics* 37(2): 234–276
- Ross, Malcolm. 2005. Pronouns as a preliminary diagnostic of grouping Papuan languages. In: *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-Speaking Peoples*, Pawley A, Attenborough R, Golson J, Hide R (eds), 15–65. Canberra: Pacific Linguistics.
- Sampson, Geoffrey, David Gil & Peter Trudgill (eds). 2009. *Language Complexity as an Evolving Variable*. Oxford: Oxford University Press.
- Schapper, Antoinette. 2009. Possession in Kemak. Paper presented at the 11th International Conference on Austronesian Linguistics, June 21–25. Aussois, France.
- Schapper, Antoinette. 2010a. Bunaq, a Papuan Language of Central Timor. PhD dissertation, The Australian National University.

- Schapper, Antoinette. 2010b. Neuter gender in eastern Indonesia. *Oceanic Linguistics* 49(2): 407–435.
- Schapper, Antoinette. 2011. Crossing the border: Historical and linguistic divides among the Bunaq in central Timor. *Wacana, Journal of the Humanities of Indonesia* 13(1): 29–49.
- Schapper, Antoinette. 2015. Wallacea, a linguistic area. *Archipel* 90: 99–151.
- Schapper, Antoinette. 2017. Farming and the Trans-New Guinea family: A consideration. In *Language dispersal beyond farming*, Martine Robbeets (ed), 155–182. Amsterdam: John Benjamins.
- Schapper, Antoinette. 2019. Build the wall! Village fortification, its timing and triggers in southern Maluku. *Indonesia and the Malay World*. Open Access URL: <https://doi.org/10.1080/13639811.2019.1554778>
- Schapper, Antoinette, Juliette Huber & Aone van Engelenhoven. 2012. The historical relation of the Papuan languages of Timor and Kisar. In *History, contact and classification of Papuan languages*, in *Language and Linguistics in Melanesia*. Harald Hammarström & Wilco van den Heuvel (eds), Special Issue 2012: 194–242. Port Moresby: Linguistic Society of Papua New Guinea.
- Schapper, Antoinette, Juliette Huber & Aone van Engelenhoven. 2014. The relatedness of Timor-Kisar and Alor-Pantar languages: a preliminary demonstration. In *Alor-Pantar Languages: History and Typology*, Marian Klamer (ed), 99–154. Berlin: Language Science Press.
- Schapper, Antoinette & Lourens de Vries. 2018. Comparatives in Melanesia: Concentric circles of convergence. *Linguistic Typology* 22(3): 437–494.
- Steinhauer, Hein. 1993. Notes on verbs in Dawanese (Timor). In *Topics in descriptive Austronesian linguistics*, Ger P. Reesink (ed), 131–158. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit te Leiden.
- Steinhauer Hein. 1996. Morphemic metathesis in Dawanese (Timor). In *Papers in Austronesian Linguistics*, Hein Steinhauer (ed), 217–232. Canberra: Pacific Linguistics.
- Steinhauer, Hein. 2014. Blagar. In *Papuan languages of Timor, Alor and Pantar. Volume 1*, Antoinette Schapper (ed), 147–218. Berlin: Mouton de Gruyter.
- Suter, Edgar. 2012. Verbs with pronominal object prefixes in Finisterre-Huon languages. *Language and Linguistics in Melanesia: Special issue on the history, contact & classification of Papuan languages*, 23–58.
- Taber, Mark. 1993. Toward a better understanding of the indigenous languages of Southwestern Maluku. *Oceanic Linguistics* 32(2): 389–441.
- Tumonggor, Meryanne K, Tatiana M Karafet, Sean Downey, J Stephen Lansing, Peter Norquest, Herawati Sudoyo, Michael F Hammer & Murray P Cox. 2014. Isolation, contact and social behavior shaped genetic diversity in West Timor. *Journal of Human Genetics* 59: 494–503
- Usher, Timothy & Antoinette Schapper. Ms. The Greater West Bomberai Family. A demonstration.
- Veloso, Alexandre. 2016. A Grammar Sketch of Naueti, a Language of East Timor. MA dissertation, Leiden University.

- Wellfelt, Emilie. 2016. *Historyscapes in Alor. Approaching Indigenous Histories in Eastern Indonesia*. PhD dissertation, Linnaeus University, Sweden.
- Willemsen, Jeroen. 2015. *Function and Grammaticalisation in Alor-Pantar Asymmetrical Verb Serialisation*. MA dissertation, Leiden University.